



## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<b>(51) International Patent Classification <sup>6</sup> :</b> <b>C07H 15/12, C12N 15/00, C12Q 1/68</b>		<b>A1</b>	<b>(11) International Publication Number:</b> <b>WO 95/28411</b>
			<b>(43) International Publication Date:</b> 26 October 1995 (26.10.95)
<b>(21) International Application Number:</b> PCT/US95/04463		<b>(81) Designated States:</b> AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).	
<b>(22) International Filing Date:</b> 12 April 1995 (12.04.95)			
<b>(30) Priority Data:</b> 08/226,972 13 April 1994 (13.04.94) US 08/404,531 15 March 1995 (15.03.95) US		<b>Published</b> <i>With international search report.</i>	
<b>(71) Applicants:</b> BAYLOR COLLEGE OF MEDICINE [US/US]; Texas Medical Center, One Baylor Plaza, Houston, TX 77030 (US). THE BOARD OF REGENTS, THE UNIVERSITY OF TEXAS SYSTEM [US/US]; 201 West 7th Street, Austin, TX 78701 (US).			
<b>(72) Inventors:</b> BRYAN, Joseph; 2605 Rice Boulevard, Houston, TX 77005 (US). BRYAN, Lydia, Aguilar; 2605 Rice Boulevard, Houston, TX 77005 (US). NELSON, Daniel, A.; 5222 McKnight Boulevard, Houston, TX 77035 (US). THOMAS, Pamela, M.; 7927 Duffield Lane, Houston, TX 77071 (US). COTE, Gilbert, J.; 1609 Morse Street, Houston, TX 77019 (US). GAGEL, Robert, F.; 6540 Vanderbilt Street, Houston, TX 77005-3823 (US).			
<b>(74) Agents:</b> CALDWELL, John, W. et al.; Woodcock Washburn Kurtz Mackiewicz & Norris, One Liberty Place - 46th Floor, Philadelphia, PA 19103 (US).			
<b>(54) Title:</b> SEQUENCE ENCODING MAMMALIAN SULFONYLUREA RECEPTOR AND METHOD OF DETECTING PERSISTENT HYPERINSULINEMIC HYPOGLYCEMIA OF INFANCY			
<b>(57) Abstract</b> <p>The present invention is directed to a method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising obtaining a sample comprising patient nucleic acids from a patient tissue sample; amplifying sulfonylurea receptor specific nucleic acids from said patient nucleic acids to produce a test fragment; obtaining a sample comprising control nucleic acids from a control tissue sample; amplifying control nucleic acids encoding wild type sulfonylurea receptor to produce a control fragment; comparing the test fragment with the control fragment to detect the presence of a sequence difference in the test fragment, wherein a difference in said test fragment indicates persistent hyperinsulinemic hypoglycemia of infancy. A diagnostic kit and primers for the detection of persistent hyperinsulinemic hypoglycemia of infancy are also within the scope of the present invention.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

**SEQUENCE ENCODING MAMMALIAN SULFONYLUREA RECEPTOR  
AND METHOD OF DETECTING PERSISTENT HYPERINSULINEMIC  
HYPOGLYCEMIA OF INFANCY**

**REFERENCE TO RELATED APPLICATIONS**

- 5           This application is a continuation-in-part of U.S. patent application Serial No. 08/226,972, filed April 13, 1994, the disclosure of which is hereby incorporated by reference in its entirety.

**REFERENCE TO GOVERNMENT GRANTS**

- 10           This work was supported in part by research grants from the National Institutes of Health, grant number NIH R01DK41898 and R01DK44311. The United States Government may have certain rights in this invention.

**BACKGROUND OF THE INVENTION**

- 15           Sulfonylureas are oral hypoglycemics widely used in the treatment of Non-Insulin Dependent Diabetes Mellitus (NIDDM). They enter the bloodstream, bind with high affinity to a pancreatic  $\beta$ -cell plasma membrane protein termed the sulfonylurea receptor, and stimulate insulin release. The  
20 mechanism of stimulation is thought to be through inhibition of an ATP-sensitive K<sup>+</sup> channel (K<sub>ATP</sub>), a key protein which sets the  $\beta$ -cell resting membrane potential (Ashcroft, et al. *Cell. Signal.* **1990**, 2, 197-214, all references cited herein are incorporated by reference in their entirety). A  
25 reduction in potassium outflow causes depolarization of the plasma membrane, activation of L-type voltage-dependent calcium channels (VDCCs), and increased cytosolic calcium. This triggers insulin release by as yet unknown mechanisms (Rajan, et al. *Diabetes Care* **1990**, 13, 340-363). In NIDDM  
30 patients on sulfonylureas, the consequent reduction in blood glucose to more normal levels is thought to be critical in controlling the disease (Gerich, J.E. *New Engl. J. Med.* **1989**, 321, 1231-1245).

- 2 -

The biochemistry of the sulfonylurea receptor (SUR) (Ashcroft et al. *Biochem. Biophys. Acta* **1992**, 1175, 45-49 and Panten et al. *Horm. Metab. Res.* **1992**, 24, 549-554) is consistent with the electrophysiology of the  $\beta$ -cell  $K_{ATP}$  channel. The endogenous regulators of channel activity include cytosolic nucleotides (ATP and Mg-ADP) and possibly phosphorylation. In the absence of cytosolic nucleotides, sulfonylureas weakly inhibit channel activity (Schwanstecher et al. *Br. J. Pharmacol* **1992**, 107, 87-94). When channels are activated by Mg-ADP, inhibition by ATP is strongly promoted by the presence of sulfonylureas. These results are interpreted as evidence that simultaneous occupancy of two nucleotide binding sites is required for effective channel inhibition by the sulfonylureas. The reported allosteric interactions correlate well with evidence that the brain receptor has two nucleotide binding sites (de Weille, et al. *J. Biol. Chem* **1992**, 267, 4557-4563) physically located on the same polypeptide chain as the sulfonylurea binding site (Bernardi et al. *Biochemistry* **1992**, 31, 6328-6332). One binding site appears to be specific for ATP, and is proposed to be the same site at which micromolar concentrations of ATP inhibit the  $K_{ATP}$  channel. A second site has high affinity for Mg-ADP, with occupancy at this site promoting channel opening. Absolute concentrations of ATP and ADP in the cell are thought to regulate channel activity in a straightforward fashion (Hopkins et al. *J. Membrane Biol.* **1992**, 129, 287-295). High ATP concentrations as a result of high serum glucose levels close the channel, stimulating insulin secretion. Reduced glucose levels increase intracellular ADP concentrations, and thereby increase the open channel probability, and decrease insulin secretion.

Although sulfonylureas, particularly tolbutamide and more potent second generation drugs like glyburide and glipizide, are considered to be relatively specific inhibitors of the  $K_{ATP}$  channel, the exact relationship between the sulfonylurea receptor and the  $K_{ATP}$  channel is not clear (Nichols et al. *Am. J. Physiol.* **1991**, 261, H1675-H1686,



Takano et al. *Progress in Neurobiology* 1993, 41, 21-30, and Edwards et al. *Annu. Rev. Pharmacol. Toxicol.* 1993, 33, 597-637). In the insulin-secreting CRI-G1 cell line, the addition of glyburide, or tolbutamide to inside-out plasma membrane patches inhibits the  $K_{ATP}$  channel (Khan et al. *Proc. R. Soc. Lond. B.* 1993, 253, 225-231), intimating direct interactions between sulfonylureas and the channel protein. In another insulin secreting cell line, CRI-D11 cells, however, the loss of sulfonylurea binding sites with the retention of  $K_{ATP}$  activity suggests these two activities may uncouple and reside on separate, transiently bound subunits (Khan et al. *Proc. R. Soc. Lond. B.* 1993, 253, 225-231). Similarly, in other cell and tissue types, sulfonylurea binding and channel activity may be uncoupled (Ashford et al. *Br. J. Pharmac.* 1990, 101, 531-540). A technique is not currently available to assess whether  $K_{ATP}$  activity resides within the same polypeptide containing the putative nucleotide and sulfonylurea binding sites, or on separate loosely, or tightly bound subunits.

A previous attempt to purify the receptor from hamster insulin-secreting tumor (HIT) cells was limited by the low abundance of the receptor and the presence of a more abundant co-purifying protein. Aguilar-Bryan, L., et al., *JBC*, 1990, 265, 8218.

The sulfonylurea receptor is the target for drugs used in the treatment of type II diabetes (non-insulin diabetes mellitus). This association has suggested it plays a role in the regulation of insulin secretion by glucose and makes the sulfonylurea receptor a potential diabetes candidate gene.

Persistent hyperinsulinemic hypoglycemia of infancy (PHHI) is an autosomal recessive disorder of glucose homeostasis characterized by unregulated secretion of insulin and profound hypoglycemia. A. Aynsley-Green et al., *Arch. Dis. Child.* 1981, 56, 496. The pathophysiology of this disease remains obscure, but *in vitro* studies suggest a defect of glucose-regulated insulin secretion in pancreatic

- 4 -

islet  $\beta$ -cells. Aynsley-Green et al., *supra.*, N. Kaiser et al., *Diabetologia* **1990**, 33, 482. The incidence of PHHI has been estimated at 1/50,000 live births in a randomly mating population. G.J. Bruining, *Curr. Opin. Pediatr.* **1990**, 2, 758. However, in a Saudi Arabian population in which 51% of births occurred to parents who were first or second cousins, the incidence has been established as 1/2675 live births. P.M. Mathew et al., *Clin. Pediatr.* **1988**, 27, 148. Recently, the PHHI gene was assigned to chromosome 11p1415.1 by linkage analysis. B. Glaser et al., *Nature Genet.* **1994**, 7, 185 and P.M. Thomas, G.J. Cote, D.M. Hallman, P.M. Mathew, *Am. J. Hum. Genet.* **1995**, 56, 416-421. Candidate genes for this disorder include those involved in the  $\beta$ -cell glucose sensing mechanism and insulin secretion. Localization of PHHI to chromosome 11p excluded previously mapped genes involved in  $\beta$ -cell function. Considered as a candidate was the newly cloned high-affinity SUR gene, a member of the ATP-binding cassette superfamily, and a putative subunit of the modulator of insulin secretion, the  $\beta$ -cell ATP-sensitive potassium channel ( $K_{ATP}$ ). S.J. Ashcroft and F. M. Ashcroft, *Biochimica et Biophysica Acta*, **1992**, 1175, 45; U. Panten, M. Schwanstecher, and C. Schwanstecher, *Horm. Metab. Res.* **1992**, 24, 549. The methods of the present invention map the PHHI and provide evidence that mutations in the sulfonylurea receptor are the cause of PHHI.

Accordingly, there remains a need to identify sulfonylurea receptor and sequences encoding sulfonylurea receptor which will provide:

1. a correlation between sulfonylurea receptor and one or more forms of diabetes,
2. a sequence to purify human sulfonylurea receptors,
3. an isolated sulfonylurea receptor, prepared by recombinant methods,

- 5 -

4. polyclonal and monoclonal antibodies and methods of preparing the same against sulfonylurea receptor,
5. information as to whether this receptor-ion channel family involves multi-subunits within each channel for channel activity,
6. gene therapy such that sequences which encode mutant sulfonylurea receptors are replaced by wild type sulfonylurea receptor sequences,
7. a method of screening to identify drugs which react with and bind to the sulfonylurea receptor,
8. non-human transgenic animals to study diabetes and PHHI, and the physiologic effects of varying levels of sulfonylurea receptor, by using an inducible promoter to regulate the expression of the sulfonylurea receptor, for example, and
9. probes, including PCR probes, for diagnosing conditions associated with the expression of a specific sulfonylurea receptor allele.

The present invention reveals that the sequence encoding the mammalian sulfonylurea receptor maps to the sequence encoding persistent hyperinsulinemic hypoglycemia of infancy.

#### SUMMARY OF THE INVENTION

The present invention provides sequences encoding a sulfonylurea receptor. Nucleic acid sequences, SEQ ID NO: 1, 4, 6, 26, 27, 32, 33, 35, and 36 are cDNA sequences to which the present invention is directed. The nucleic acid sequences, SEQ ID NOS: 2 and 3, may be added to the 5' end of SEQ ID NO: 1 to furnish a sequence encoding a sulfonylurea receptor. SEQ ID NOS: 32, 33, 35 and 36 are rodent sequences (SEQ ID NOS: 32 and 33 - rat, SEQ ID NOS: 34 and 35 - hamster) encoding sulfonylurea receptor which functionally bind sulfonylurea. SEQ ID NOS: 26 and 27 are human sequences which encode sulfonylurea receptor. SEQ ID NOS: 27, 33, and 36 set forth DNA sequences translated into amino acid sequences, which set forth below the DNA sequence.

- 6 -

A further aspect of the present invention provides sulfonylurea receptor polypeptides and/or proteins. SEQ ID NOS: 4, 5, 6, 7, 26, and 28 are novel polypeptides of the invention produced from nucleotide sequences encoding rat (SEQ ID NOS: 4, 5, and 34), hamster (SEQ ID NOS: 6, 7, and 37), and human (SEQ ID NOS: 26 and 28) sulfonylurea receptor, respectively. Also within the scope of the present invention is a purified sulfonylurea receptor.

The present invention also provides nucleic acid sequences encoding a sulfonylurea receptor, expression vectors comprising a nucleic acid sequence encoding a sulfonylurea receptor, transformed host cells capable of expressing a nucleic acid sequence encoding a sulfonylurea receptor, cell cultures capable of expressing a sulfonylurea receptor, and protein preparations comprising a sulfonylurea receptor.

A method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising obtaining a sample comprising patient nucleic acids from a patient tissue sample; amplifying sulfonylurea receptor specific nucleic acids from said patient nucleic acids to produce a test fragment; obtaining a sample comprising control nucleic acids from a control tissue sample; amplifying control nucleic acids encoding wild type sulfonylurea receptor to produce a control fragment; comparing the test fragment with the control fragment to detect the presence of a sequence difference in the test fragment, wherein a difference in said test fragment indicates persistent hyperinsulinemic hypoglycemia of infancy is also an embodiment of the present invention.

Other methods of the present invention include a method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising obtaining a sample comprising patient mRNA from a patient tissue sample; reverse transcribing said mRNA into cDNA to produce patient cDNA; amplifying sulfonylurea receptor specific cDNA from said patient cDNA to produce amplified patient cDNA; obtaining a sample comprising

- 7 -

control nucleic acids from a control tissue sample; amplifying control DNA encoding wild type sulfonylurea receptor to produce control cDNA; digesting said test fragment and said control fragment with a selected  
5 endonuclease; and comparing the test fragment to the control fragment, wherein said test fragment indicates persistent hyperinsulinemic hypoglycemia of infancy.

Another embodiment of the present invention is a diagnostic kit for detecting persistent hyperinsulinemic  
10 hypoglycemia of infancy comprising in one or more containers a pair of primers, wherein one primer within said pair is complementary to a region of the sulfonylurea receptor, wherein one of said pair of primers is selected from the group consisting of SEQ ID NOS: 16-24, a probe specific to  
15 the amplified product, and a means for visualizing amplified DNA, such as and not limited to fluorescent stain,  $^{32}\text{P}$ , and biotin, and optionally including one or more size markers, positive and negative controls, and restriction endonucleases.

20 Still another embodiment of the present invention includes the primer sequences identified in SEQ ID NOS: 16-24.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 displays characteristics of the purified  
25 HIT cell receptor. The radiolabeled receptor (lanes 1 and 3) cleaved with endoglycosidase F/N-glycosidase F (endo F), increases the mobility of the protein by approximately 3 kDa (lane 2). Subsequent partial V8 protease digestion (lanes 4 and 6) yielded radiolabeled fragments that also shift  
30 mobility with endo F treatment (lane 5). Each of these species has the same N-terminal sequence (left side of figure), except that receptor deglycosylation results in an Asp at residue 9.

Figure 2A shows that antibodies against residues  
35 10-20 specifically recognize the 140 kDa polypeptide. Purified 140 kDa polypeptide was electrophoresed on a single

- 8 -

lane of a 6% SDS gel and transferred to Immobilon P. The Immobilon P was placed in a miniblottedter and the lanes incubated as follows: Lane 1 - Preimmune serum. Lane 2 - Immune serum. Lane 3 - Immune serum + immunogen. Lane 4 -  
5 immune serum + irrelevant MAP peptide. The filter was further incubated with a second antibody (goat anti-rabbit conjugated to alkaline phosphatase) and developed with the appropriate substrates. Figure 2B displays antibodies which recognize a polypeptide with the appropriate mobility shift  
10 following Endo F treatment. Purified receptor (lane 1) was incubated for 30 min at 37 °C in the presence (lane 2), or absence (lane 3) of endoglycosidase F/N-glycosidase F, incubated with first (anti-MAP 10-20) and second antibody, and developed with substrate. The bottom panel shows the  
15 autoradiogram of the immunoblot in the top panel. Figure 2C shows antibodies which immunoprecipitate the photolabeled 140 kDa receptor. HIT cell membranes were incubated with <sup>125</sup>I-labeled iodoglyburide, photolabeled, solubilized with 1% digitonin, centrifuged at 100,000 x g and the supernatant  
20 (lane 1) incubated with preimmune serum (lane 2), immune serum (lane 3), immune serum + anti-MAP 10-20 (lane 4) and immune serum + irrelevant MAP peptide (lane 5). Samples were co-incubated with protein A-Sepharose, the beads washed with buffer, heated in the presence of pH 9 sample buffer,  
25 electrophoresed on a 6% polyacrylamide SDS gel, and an autoradiogram prepared. Results using antibodies against receptor residues 1-8 were the same as those using antibodies against residues 10-20.

Multiple antigenic peptides (MAPS) were synthesized  
30 (Posnett et al. *J. Biol. Chem.* 1988 263:1719-1725) and polyclonal antibodies generated in rabbits produced by standard methods (Antibodies: A Laboratory Manual Cold Spring Harbor Laboratory 1988). Interdermal injections of 1 mg of antigen were spaced 2-3 weeks apart, and contained complete,  
35 or incomplete Freund's adjuvant

Figure 3 shows the nucleotide sequence of cDNAs encoding the high affinity sulfonylurea receptors from rat

- 9 -

and hamster together with the deduced amino acid sequences. DNA sequencing was performed by the dideoxy chain termination method with <sup>35</sup>S-labeled nucleotides. Numbers on the left of each column indicate the nucleotide positions; numbers on the right of each column indicate the amino acid positions. The amino acid sequence of four peptides determined by chemical methods are marked in bold italics. The putative ATP-binding domains are shown underlined and the Walker (Walker et al. *EMBO Jour.* **1982** 1:945-951) "A" and "B" consensus sequences are boxed in dotted lines. The Walker A sequence, GXXGXGK (SEQ ID NO: 14) and B sequence, X<sup>1</sup>X<sup>1</sup>X<sup>1</sup>D (SEQ ID NO: 15) (where X<sup>1</sup> is a hydrophobic amino acid) are part of a common nucleotide binding fold found in a number of nucleotide binding proteins. The G residues in the A sequence interact with a phosphate in the nucleotide. Possible sites of phosphorylation by protein kinase A (closed diamonds) and C (open diamonds) are indicated. Potential transmembrane spanning helices assigned using the algorithm of Eisenberg et al. (Eisenberg, et al. *J. Mol. Biol.* **1984** 179:125) are boxed.

The nucleotide sequence and first row of amino acid sequence are for the rat; the amino acid sequence differences are given in the second row for the hamster. A consensus poly A addition site is indicated in bold at the 3' end of the cDNA. Abbreviations for the amino acid residues are: A, ala; C, cys; D, asp; E, glu; F, phe; G, gly; H, his; I, ile; K, lys; L, leu; M, met; N, asn; P, pro; Q, gln; R, arg; S, ser, T, thr; V, val; W, trp; and Y, tyr.

Figure 4 displays the deduced amino acid sequence of the sulfonylurea receptor aligned with the amino acid sequence of a MRP, multidrug resistance-associated protein, (dvhuar). The alignment was generated with PILEUP from the Genetics Computer Group package (Madison, WI) (Genetics Computer Group Program Manual for the GCG Package Version 7 **1991**) using a modification of the algorithm of Feng and Doolittle (Feng, et al. *J. Mol. Evol.* **1987** 25:351).

Figure 5 exhibits the sequence comparison of nucleotide binding folds. The alignment was generated using

- 10 -

the PILEUP program as described in Figure 4. The Walker A and B consensus sites are given in boldface. Upper case letters indicate a match to the consensus sequence generated using the pretty program (Genetics Computer Group, Madison, WI).

Figure 6 is a northern blot of total RNA from  $\alpha$ - and  $\beta$ -cell lines hybridized with a 2.2 kb EcoRI-XhoI fragment of the sulfonylurea receptor. Approximately 10  $\mu$ g of RNA from (A)  $\alpha$ TC-6 cells, (B) HIT cells, (C) RIN cells and (D) mouse liver was analyzed using standard procedures (Ausubel et al. *Current Protocols in Mol. Biol.* 1994). The estimated size of the major component is approximately 5000 nucleotides.

Figure 7 displays a hydrophobicity profile of the Rat Sulfonylurea receptor. Hydrophobicity values were determined according to Kyte and Doolittle (Kyte et al. *J. Mol. Biol.* 1982 157:105-132) for 11-residue peptides and are plotted versus the amino acid number. The bars marked A and B are over the Walker A and B consensus sequences Walker et al. *EMBO Jour.* 1982 1:945-951).

Figure 8 shows a schematic model of the high affinity sulfonylurea receptor. The Walker A and B sites are marked within the two nucleotide binding folds. Based on the hydrophobicity and hydrophobic moment data there are nine transmembrane spanning domains before the first nucleotide binding fold and four transmembrane spanning domains between the two folds. The branched structure at the N-terminus of the mature receptor symbolizes glycosylation.

Figure 9A reveals the results of *in vitro* translation of mRNA transcribed from the rat sulfonylurea cDNA. The cDNA was subcloned into pGEM4 (Promega, Inc., Madison, WI) and transcribed using the SP6 promoter and SP6 RNA polymerase following the manufacturer's directions. RNA was translated in rabbit reticulocyte lysate (Promega, Inc.) following the manufacturer's recommendations for  $^{35}$ S-methionine. Lane 1 is the HIT cell photolabeled receptor as a marker, lane 2 is the *in vitro* translation product



- 11 -

resulting from addition of receptor mRNA and lane 3 is the result of no added RNA. The arrow marks the 140 kDa protein.

Figure 9B displays a gel of the results of immunoprecipitation of the RIN cell sulfonylurea receptors with polyclonal antibodies directed against a nucleotide binding fold domain (NBF). Lane 1; 140 and 150 kDa receptors from soluble RIN cell membrane proteins, lane 2; immunoprecipitation with preimmune serum, lane 3; immune serum from rabbit immunized with NBF2, lane 4; immune serum + NBF2 fusion protein. Sulfonylurea receptor cDNA regions encoding the NBF2 domain was subcloned in frame into pMALc and expression of the proteins fused with maltose binding protein induced in *E. coli*. Fusion proteins were purified by electrophoresis and electroelution, and 200 $\mu$ g amounts, with complete, or incomplete Freund's adjuvant, injected interdermally into rabbits using a standard 2-3 week regimen of bleeding and boosting.

Figure 10 displays the genomic organization and cDNA sequence of the human sulfonylurea receptor (SUR) homologue in the second nucleotide binding fragment region (NBF-2). The sequence encoding NBF-2 is located within SEQ ID NO: 26, nucleic acid positions 524 to 1048. Solid rectangles represent exons which are labeled  $\alpha$ - $\phi$  for identification. The numbers between rectangles represent intronic sizes. Primers used in mutational analysis are diagrammed and listed in SEQ ID NOS: 16-24.

Figure 11A-D display the exon mutation in the SUR NBF-2. Figures 11A is a schematic representation of NBF-2 exons  $\beta$ , X,  $\delta$  illustrating the normal (upper) and mutant (lower) RNA splicing patterns. Figure 11B displays the sequence of a pancreatic cDNA product, from an affected child of Family 6, demonstrating the exon skipping event. Skipping of exon X results in a 109 bp deletion in the mRNA transcript, a frame shift and inclusion of a premature stop codon. Single upper case letters represent amino acids. Figure 11C shows the sequence of genomic DNA from the affected patient in Figure 11B which reveals a G to A point

- 12 -

mutation at the 3' end of the exon, which exon is excised in mRNA, as compared to a normal sample of genomic DNA. Exonic sequence is in upper case and intronic sequence in lower case letters. Figure 11D shows *MspI* restriction enzyme analysis of PCR-amplified genomic DNA from members of Family 6, indicating affected individuals. The G to A mutation destroyed a restriction site for *MspI* (C/CGG). Normal PCR product is digested into 304 bp, 85 bp, and 38 bp fragments, while that containing the mutation is digested into 304 bp and 123 bp fragments. MW is 100 bp ladder (GIBCO-BRL, Gaithersburg, Maryland), UC is an uncut sample, C is a control PCR reaction lacking template.

Figure 12A-D reveal a mutation in the intron preceding NBF-2 exon  $\alpha$ , which activates cryptic 3' splice site usage. Figure 12A displays the sequence of genomic DNA from an affected member of Family 4 which revealed a G to A mutation in the splice site preceding the first exon of the NBF-2. Figure 12B shows *NciI* restriction enzyme analysis of genomic DNA from members of Family 4, indicating affected individuals. The G to A mutation destroys a restriction site for *NciI* (CC/(G/C)GG). Normal PCR product is digested into 71 bp and 75 bp fragments, while that containing the mutant sequence is not cut. By previous haplotype analysis, the unaffected sibling in this family had two wild type alleles, P.M. Thomas, G.J. Cote, D.M. Hallman, P.M. Mathew, *Am. J. Hum. Genet.*, *supra*. Figure 12C illustrates the constructs used to examine RNA processing of exons within NBF-2. Solid rectangles and thin lines represent human *SUR* gene exonic and intronic sequences, respectively. The unmarked solid rectangle represents a portion of the exon which is 5' to exon  $\alpha$  of the NBF-2 region. That labeled RSV represents the enhancer and promoter isolated from the rous sarcoma virus long terminal repeat. The thick line represents an intronic sequence derived from vector and the human metallothioneine IIA gene, which also contains polyadenylation signals. Normal and mutant RNA splicing patterns, including the location of the three cryptic splice sites, are diagrammed in

- 13 -

the lower portion. The open triangle marks the position of the mutated base within the splice site. Figure 12D shows PCR amplification across splice site of normal (N) and mutant (M) cDNA transcripts, isolated 48 hours after transfection with the splicing constructs. Subcloning and sequencing of these products revealed their identity as diagrammed in Figure 12C. The control (C) represents cDNA amplified from untransfected cells.

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to the nucleic acid and protein sequences encoding a sulfonylurea receptor. The present invention provides a nucleotide sequence of a sulfonylurea receptor, Figure 3, and SEQ ID NOS: 1, 4, 6, 33, and 36. Novel polypeptide sequences, SEQ ID NOS: 4, 5, 6, 7, 34, and 37 coding for a sulfonylurea receptor are also included in the present invention. SEQ ID NOS: 26, 27, and 28 provide the nucleic acid and amino acid sequences of the last 11 exons of the 3' end of human sulfonylurea receptor, hereinafter referred to, together with the rodent sequences for sulfonylurea receptor, as sequence for the sulfonylurea receptor.

SEQ ID NO: 1 provides the cDNA sequence of a rodent sulfonylurea receptor. SEQ ID NOS: 26 and 27 provide the human genomic DNA sequence of sulfonylurea receptor. Nucleic acids within in the scope of the present invention include cDNA, RNA, genomic DNA, sequences within these larger sequences, antisense oligonucleotides. Sequences encoding the sulfonylurea receptor also include amino acid, polypeptide, and protein sequences. Variations in the nucleic acid and polypeptide sequences of the present invention are within the scope of the present invention and include N terminal and C terminal extensions, transcription and translation modifications, and modifications in the cDNA sequence to facilitate and improve transcription and translation efficiency. In addition, mismatches within the sequences identified herein, which achieve the methods of the

- 14 -

invention, such that the mismatched sequences are substantially complementary to the sulfonylurea receptor sequences identified, are also considered within the scope of the present invention. Mismatches which permit substantial  
5 complementarity to the sulfonylurea receptor sequences, such as similarity in residues in hydrophobicity, will be known to those of skill in the art once armed with the present disclosure. In addition, the sequences of the present invention may be natural or synthetic.

10 SEQ ID NOS: 2 and 3 are rat and hamster sequences, respectively, which may be added to SEQ ID NO: 1 at the 5' end preceding the coding region of the sulfonylurea gene.

The cDNA and amino acid sequences identified in Figure 3 indicate several amino acid residue differences  
15 between rat and hamster sulfonylurea receptor. These differences are indicated in the corresponding nucleic acid sequence of SEQ ID NO: 1 by an "N" in the positions which result in different amino acids. Accordingly, "N" appears in SEQ ID NO: 1, at nucleic acid positions 586, 752, 802, 842,  
20 1204, 1318, 1518, 1711, 2000, 2009, 2245, 2275, 2278, 2280, 2281, 2288, 2305, 2383, 2384, 2518, 2540, 2548, 2848, 2850, 2913, 3173, 3208, 3230, 3463, 3601, 3751, 4442. "N" may be any nucleic acid residue which provides an amino acid which results in a sulfonylurea receptor. For example, SEQ ID NOS:  
25 4 and 6 provide cDNAs of rat and hamster, respectively, where N of SEQ ID NO: 1 is identified as a particular nucleic acid such that a corresponding amino acid may be provided in the protein. In addition, an Asparagine is inserted in the hamster sequence at amino acid position 740 and a threonine  
30 is deleted in the hamster sequence at amino acid position 831. The N positions indicated for SEQ ID NO: 1 result in differences in the rat and hamster sequences provided in SEQ ID NOS: 4 and 5 (rat), 6 and 7 (hamster) at amino acid positions 196 - Val, Ile; 251 - Gly, Ala; 268 - Leu, Val; 281  
35 - Gln, Pro; 300 - Val, Ile; 402 - Leu, Met; 440 - Ala, Thr; 491 - Tyr, His; 506 - Ile, Met; 516 - Asn, Ser; 524 - Lys, Val; 571 - Phe, Leu; 667 - Thr, Ala; 670 - Thr, Met; 749 -

- 15 -

Arg, 750 - Gly; 759 - Gly, 760 - Ser; 760 - Arg, 761 - Trp;  
761 - Phe, 762 - Leu; 763 - Cys, 764 - Tyr; 769 - Cys, 770 -  
Arg; 795 - Lys, 796 - Pro; (two series of numbers appear due  
to the asparagine deletion in the hamster sequence, the  
5 hamster codon is represented by the second number) 840 - Arg,  
Gly; 847 - Thr, Arg; 850 - Ile, Phe; 950 - Pro, Ser; 971 -  
Asp, Glu; 1058 - Ala, Asp; 1070 - Val, Leu; 1077 - Ala, Val;  
1155 - Ala, Thr; 1201 - Leu, Val; 1251 - Arg, Cys; and 1481 -  
Lys, Thr. In addition, the hamster sequence has an Asn  
10 insert at position 741 and a Thr deletion at position 831.

A purified sulfonylurea receptor is also provided  
by the present invention. The purified sulfonylurea receptor  
may have an amino acid sequence as provided by SEQ ID NOS: 4,  
5, 6, 7, 34, and 37. A purified sulfonylurea receptor having  
15 the amino acid sequence of SEQ ID NO: 28 is also within the  
scope of the present invention.

The present invention is directed to sulfonylurea  
receptor sequences obtained from mammals from the Order  
Rodentia, including and not limited to hamsters, rats, and  
20 mice; Order Logomorpha, such as rabbits; more particularly  
the Order Carnivora, including Felines (cats) and Canines  
(dogs); even more particularly the Order Artiodactyla,  
Bovines (cows) and Suines (pigs); and the Order  
Perissodactyla, including Equines (horses); and most  
25 particularly the Order Primates, Ceboids and Simoids  
(monkeys) and Anthropoids (humans and apes). The mammals of  
most preferred embodiments are humans.

There are several transfection techniques by which  
a sulfonylurea receptor may be obtained. An appropriate RNA  
30 may be hybridized to a cDNA to obtain a sulfonylurea receptor  
nucleic acid sequence. A nucleic acid sequence encoding  
sulfonylurea receptor may be inserted into cells and the  
corresponding protein immunoprecipitated with an antibody.  
Labeled drugs known to bind sulfonylurea receptor protein may  
35 be added to cell culture to label the receptor. The drug  
labeling procedure may involve modifying cells such that the  
cell culture provides conditions similar to  $\beta$  cells, cells

- 16 -

where sulfonylurea receptors naturally appear; and the sulfonylurea receptor may be part of a larger multisubunit ATP receptor channel, which may not be provided by the cells in culture.

5           Generally, the sequences of the invention may be produced in host cells transformed with an expression vector comprising a nucleic acid sequence encoding the sulfonylurea receptor. The transformed cells are cultured under conditions whereby the nucleic acid sequence coding for the  
10 sulfonylurea receptor is expressed. After a suitable amount of time for the protein to accumulate, the protein is purified from the transformed cells.

A gene coding for sulfonylurea receptor may be obtained from a cDNA library. Suitable libraries can be  
15 obtained from commercial sources such as Clontech, Palo Alto, CA. Libraries may also be prepared using the following non-limiting examples hamster insulin-secreting tumor (HIT), mouse  $\alpha$ TC-6, and rat insulinoma (RIN) cells. Positive clones are then subjected to DNA sequencing to determine the  
20 presence of a DNA sequence coding for sulfonylurea receptor. DNA sequencing is accomplished using the chain termination method of Sanger et al., *Proc. Nat'l. Acad. Sci, U.S.A.*, 1977, 74, 5463. The DNA sequence encoding sulfonylurea receptor is then inserted into an expression vector for later  
25 expression in a host cell.

Expression vectors and host cells are selected to form an expression system capable of synthesizing sulfonylurea receptor. Vectors including and not limited to baculovirus vectors may be used in the present invention.  
30 Host cells suitable for use in the invention include prokaryotic and eukaryotic cells that can be transformed to stably contain and express sulfonylurea receptor. For example, nucleic acid coding for the recombinant protein may be expressed in prokaryotic or eukaryotic host cells,  
35 including the most commonly used bacterial host cell for the production of recombinant proteins, *E. coli*. Other microbial strains may also be used, however, such as *Bacillus subtilis*,

- 17 -

and other enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcescens*, various species of *Pseudomonas*, or other bacterial strains.

Commonly used eukaryotic systems include yeast,  
5 such as *Saccharomyces cerevisiae*; insect cells, such as *Spodoptera frugiperda*; chicken cells, such as E3C/O and SL-29; mammalian cells, such as HeLa, Chinese hamster ovary cells (CHO), COS-7 or MDCK cells and the like. The foregoing list is illustrative only and is not intended in any way to  
10 limit the types of host cells suitable for expression of the nucleic acid sequences of the invention.

As used herein, expression vectors refer to any type of vector that can be manipulated to contain a nucleic acid sequence coding for sulfonylurea receptor, such as  
15 plasmid expression vectors and viral vectors. The selection of the expression vector is based on compatibility with the desired host cell such that expression of the nucleic acid encoding sulfonylurea receptor results. Plasmid expression vectors comprise a nucleic acid sequence of the invention  
20 operably linked with at least one expression control element such as a promoter. In general, plasmid vectors contain replicon and control sequences derived from species compatible with the host cell. To facilitate selection of plasmids containing nucleic acid sequences of the invention,  
25 plasmid vectors may also contain a selectable marker such as a gene coding for antibiotic resistance. Suitable examples include the genes coding for ampicillin, tetracycline, chloramphenicol or kanamycin resistance.

Suitable expression vectors, promoters, enhancers,  
30 and other expression control elements are known in the art and may be found in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, second edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989), incorporated herein by reference in its entirety.

35 Transformed host cells containing a DNA sequence encoding sulfonylurea receptor may then be grown in an appropriate medium for the host. The cells are then grown

- 18 -

until product accumulation reaches desired levels at which time the cells are then harvested and the protein product purified in accordance with conventional techniques.

Suitable purification methods include, but are not limited

- 5 to, SDS PAGE electrophoresis, phenylboronate-agarose, reactive green 19-agarose, concanavalin A sepharose, ion exchange chromatography, affinity chromatography, electrophoresis, dialysis and other methods of purification known in the art.

- 10 Protein preparations, of purified or unpurified sulfonylurea receptor produced by host cells, are accordingly produced which comprise sulfonylurea receptor and other material such as host cell components and/or cell medium, depending on the degree of purification of the protein.

- 15 Antibodies, including and not limited to monoclonal, polyclonal, and chimeric, prepared and used against a sulfonylurea receptor are also within the scope of the present invention, and may be prepared by methods known to those of skill in the art such as and not limited to the  
20 methods of Kohler and Milstein, Nature, 256: 495-497 (1975), incorporated herein by reference in its entirety.

- The invention also includes a transgenic non-human animal, including and not limited to mammals, such as and not limited to a mouse, rat, or hamster, whose germ cells and  
25 somatic cells contain a sequence encoding a sulfonylurea receptor introduced into the animal or an ancestor of the animal. The sequence may be wild-type or mutant and may be introduced into the animal at the embryonic or adult stage. The sequence is incorporated into the genome of an animal  
30 such that it is chromosomally incorporated into an activated state. Embryo cells may be transfected with the gene as it occurs naturally, and transgenic animals are selected in which the gene has integrated into the chromosome at a locus which results in activation. Other activation methods  
35 include modifying the gene or its control sequences prior to introduction into the embryo. The embryo may be transfected using a vector containing the gene.



- 19 -

In addition, a transgenic non-human animal may be engineered wherein the sulfonylurea receptor is suppressed. For purposes of the present invention, suppression of the sulfonylurea receptor includes, and is not limited to

5 strategies which cause the sulfonylurea receptor not to be expressed. Such strategies may include and are not limited to inhibition of protein synthesis, pre-mRNA processing, or DNA replication. Each of the above strategies may be accomplished by antisense inhibition of sulfonylurea receptor

10 gene expression. Many techniques for transferring antisense sequences into cells are known to those of skill, including and not limited to microinjection, viral-mediated transfer, somatic cell transformation, transgene integration, and the like, as set forth in Pinkert, Carl, *Transgenic Animal*

15 *Technology*, 1994, Academic Press, Inc., San Diego, CA, incorporated herein by reference in its entirety.

Further, a transgenic non-human animal may be prepared such that the sulfonylurea receptor gene is knocked out. For purposes of the present invention, a knock out

20 includes and is not limited to disruption or rendering null the sulfonylurea receptor gene. A knock out may be accomplished, for example, with antisense sequences for the sulfonylurea receptor mutating the sequence for the sulfonylurea receptor. The sulfonylurea receptor gene may be

25 knocked out by injection of an antisense sequence for all or part of the sulfonylurea receptor sequence such as an antisense sequence for all or part of SEQ ID NO: 27. Once the sulfonylurea receptor has been rendered null, correlation of the sulfonylurea receptor to persistent hyperinsulinemic

30 hypoglycemia of infancy may be tested. Sequences encoding mutations affecting the sulfonylurea receptor may be inserted to test alterations in glucose homeostasis.

Also in transgenic non-human animals, the sulfonylurea receptor may be replaced by preparing a

35 construct having an insulin promoter ligated to the sulfonylurea receptor gene. This experiment permits testing

- 20 -

of mutant sulfonylurea receptors directly in the pancreas of the transgenic animal.

Transgenic non-human animals may also be useful for testing nucleic acid changes to identify nucleotides which are responsible for ADP and ATP modulation of the sulfonylurea receptor resulting in an increase or decrease in glucose sensitivity of insulin release.

The present invention is also directed to gene therapy wherein a mutant sulfonylurea receptor is replaced by a wild type sulfonylurea receptor. A resulting transgenic non-human animal thus comprises a recombinant sulfonylurea receptor. In addition, gene therapy techniques may be used for individuals with persistent hyperinsulinemic hypoglycemia of infancy. For purposes of the present invention, gene therapy refers to the transfer and stable insertion of new genetic information into cells for the therapeutic treatment of diseases or disorders. The foreign gene is transferred into a cell that proliferates to spread the new gene throughout the cell population. Known methods of gene transfer include microinjection, electroporation, liposomes, chromosome transfer, transfection techniques, calcium-precipitation transfection techniques, and the like.

Numerous techniques are known in the art for the introduction of foreign genes into cells and may be used to construct the recombinant cells for purposes of gene therapy, in accordance with this embodiment of the invention. The technique used should provide for the stable transfer of the heterologous gene sequence to the stem cell, so that the heterologous gene sequence is heritable and expressible by stem cell progeny, and so that the necessary development and physiological functions of the recipient cells are not disrupted. Techniques which may be used include but are not limited to chromosome transfer (e.g., cell fusion, chromosome-mediated gene transfer, micro cell-mediated gene transfer), physical methods (e.g., transfection, spheroplast fusion, microinjection, electroporation, liposome carrier), viral vector transfer (e.g., recombinant DNA viruses,

- 21 -

recombinant RNA viruses) and the like (described in Cline, M. J., 1985, Pharmac. Ther. 29:69-92, incorporated herein by reference in its entirety).

5 The term "purified", when used to describe the state of nucleic acid sequences of the invention, refers to nucleic acid sequences substantially free of nucleic acid not coding for sulfonylurea receptor or other materials normally associated with nucleic acid in non-recombinant cells, i.e., in its "native state."

10 The term "purified" or "in purified form" when used to describe the state of a sulfonylurea receptor, protein, polypeptide, or amino acid sequence, refers to sulfonylurea receptor sequences free, to at least some degree, of cellular material or other material normally associated with it in its  
15 native state. Preferably the sequence has a purity (homogeneity) of at least about 25% to about 100%. More preferably the purity is at least about 50%.

To begin to elucidate the relationship between the sulfonylurea receptor and  $K_{ATP}$ , the iodinated derivative of  
20 glyburide was used to identify, and subsequently to purify and obtain N-terminal amino acid sequence from the 140 kDa high affinity, hamster insulin-secreting tumor (HIT) cell sulfonylurea receptor. The peptide sequence data was used to clone full length cDNAs encoding the rat and hamster  $\beta$ -cell  
25 proteins of the present invention.

Another embodiment of the present invention is a method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising obtaining a sample comprising patient nucleic acids from a patient tissue sample; amplifying  
30 sulfonylurea receptor specific nucleic acids from said patient nucleic acids to produce a test fragment; obtaining a sample comprising control nucleic acids from a control tissue sample; amplifying control nucleic acids encoding wild type sulfonylurea receptor to produce a control fragment;  
35 comparing the test fragment with the control fragment to detect the presence of a sequence difference in the test fragment, wherein a difference in said test fragment

- 22 -

indicates persistent hyperinsulinemic hypoglycemia of infancy is also an embodiment of the present invention.

- Persistent hyperinsulinemic hypoglycemia of infancy (PHHI) is an autosomal recessive disorder which results in
- 5 unregulated insulin secretion. The present invention revealed several different mutations in the sulfonylurea receptor in individuals with PHHI. These mutations include nucleic acid transition and restriction fragment length polymorphism, both defined herein as sequence differences.
- 10 The nucleic acid sequence transition may be a G to A transition at nucleic acid position 750 in SEQ ID NO: 26 which results in PHHI. This transition was found to occur in nine affected children in nine different families of the families studied. The pancreatic cDNA from a child with this
- 15 transition involved skipping an exon. Exon X of Figure 10 was skipped resulting in an mRNA transcript having a 109 bp deletion, a frame shift, and the inclusion of a premature stop codon. This deletion may be seen by performing rtPCR on the child's mRNA. Amplification of SEQ ID NO: 26 resulted in
- 20 a 427 base pair product for the normal as well as for the mutant genomic DNA. Digesting the normal and mutant products with *MspI*, however, resulted in three fragments (304 bp, 85 bp, and 38 bp) for the normal gene and two fragments (304 bp and 123 bp) for the mutant gene of affected children.
- 25 Another mutation involves a G to A transition in intron 11 of the human sulfonylurea receptor which gives rise to PHHI. The transition site corresponds to position 27 of SEQ ID NO: 29. The G to A transition destroys a restriction site for *NciI*. Both normal and mutant PCR products resulted
- 30 in 146 bp. Digestion with *NciI* resulted in two fragments (71 bp and 75 bp) fragments for normal individuals, while the mutant sequence was not be cut by *NciI* and thus remained at 146 bp.

- A method of detecting persistent hyperinsulinemic
- 35 hypoglycemia of infancy comprising obtaining a sample comprising patient genomic DNA from a patient tissue sample;

- 23 -

amplifying sulfonylurea receptor specific DNA from said patient genomic DNA to produce a test fragment; obtaining a sample comprising control nucleic acids from a control tissue sample; amplifying control DNA encoding wild type

5 sulfonylurea receptor to produce a control fragment; comparing the test fragment with the control fragment to detect a test fragment having G to A transition at nucleic acid position 750 of SEQ ID NO: 26, or a G to A transition at nucleic acid position 27 of SEQ ID NO: 29, wherein said test

10 fragment indicates persistent hyperinsulinemic hypoglycemia of infancy is also an embodiment of the present invention.

Also within the scope of the present invention is a method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising obtaining a sample comprising patient

15 mRNA from a patient tissue sample; reverse transcribing said mRNA into cDNA to produce patient cDNA; amplifying sulfonylurea receptor specific cDNA from said patient cDNA to produce a test fragment; obtaining a sample comprising control nucleic acids from a control tissue sample; amplifying

20 control DNA encoding wild type sulfonylurea receptor to produce a control fragment; digesting said test fragment and said control fragment with an endonuclease selected from the group consisting of *NciI* and *MspI*; and comparing the test fragment with the control fragment to detect a restriction

25 fragment length polymorphism, wherein said restriction fragment length polymorphism indicates persistent hyperinsulinemic hypoglycemia of infancy.

The restriction fragment polymorphisms include test fragments of about 304 bp and about 123 bp as a result of

30 *MspI* restriction and a test fragment of about 146 bp as a result of *NciI* restriction. The test fragments thus indicate persistent hyperinsulinemic hypoglycemia of infancy.

In accordance with methods of the present invention, methods of detecting PHHI in a patient are

35 provided comprising obtaining a patient tissue sample for testing. The tissue sample may be solid or liquid, a body fluid sample such as and not limited to blood, serum, saliva,

- 24 -

sputum, mucus, bone marrow, urine, lymph, and a tear; and feces. In addition, a tissue sample such as pancreatic tissue may be provided for the detection of PHHI in accordance with the present invention.

5 A test fragment is defined herein as an amplified sample comprising sulfonylurea receptor specific nucleic acids from a patient suspected of having PHHI. A control fragment is an amplified sample comprising normal or wild type sulfonylurea receptor specific nucleic acids from an individual not suspected of having PHHI.

The method of amplifying nucleic acids may be the polymerase chain reaction using a pair of primers wherein at least one primer within the pair is selected from the group consisting of SEQ ID NO: 16-24. When the polymerase chain reaction is the amplification method of choice, a pair of primers may be used such that one primer of the pair is selected from the group consisting of SEQ ID NOS: 17, 18, 21, and 23 and the second primer of the pair is selected from the group consisting of SEQ ID NOS: 16, 19, 20, 22, and 24.

20 Nucleic acids, such as DNA (such as and not limited to genomic DNA and cDNA) and/or RNA (such as and not limited to mRNA), are obtained from the patient sample. Preferably RNA is obtained. A whole blood gradient may be performed to isolate nucleated cells and total RNA is extracted such as by the RNazole B method (Tel-Test Inc., Friendswood, Texas) or by modification of any methods known in the art such as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 1989, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, incorporated herein by reference in its entirety.

30 Nucleic acid extraction is followed by amplification of the same by any technique known in the art. The amplification step includes the use of at least one primer sequence which is complementary to a portion of sulfonylurea receptor specific expressed nucleic acids or sequences. Primer sequences useful in the amplification methods include and are not limited to SEQ ID NOS: 16-24, which may be used in the amplification methods. Any primer

- 25 -

sequence of about 10 nucleotides to about 35 nucleotides, more preferably about 15 nucleotides to about 30 nucleotides, even more preferably about 17 nucleotides to about 25 nucleotides may be useful in the amplification step of the methods of the present invention. In addition, mismatches within the sequences identified above, which achieve the methods of the invention, such that the mismatched sequences are substantially complementary and thus hybridizable to the sequence sought to be identified, are also considered within the scope of the disclosure. Mismatches which permit substantial similarity to SEQ ID NOS: 16-24, such as and not limited to sequences with similar hydrophobicity and hydrophilicity, will be known to those of skill in the art once armed with the present disclosure. The primers may also be unmodified or modified. Primers may be prepared by any method known in the art such as by standard phosphoramidite chemistry. See Sambrook et al., *supra*.

The method of amplifying nucleic acids may be the polymerase chain reaction using a pair of primers wherein at least one primer within the pair is selected from the group consisting of SEQ ID NO: 16-24. When the polymerase chain reaction is the amplification method of choice, a pair of primers may be used such that one primer of the pair is selected from the group consisting of SEQ ID NOS: 16-24.

Primers used in mutational analysis were SEQ ID NO: 16: CACGCTCAGGTTCTGGAT; SEQ ID NO: 17: TCAACTGGATGGTGAGGA; SEQ ID NO: 18: 5' TGACATCGCCAACTGC; SEQ ID NO: 19: TCCTGGCAGTGCCTTCA; SEQ ID NO: 20: TCCTCTCAGGGTCCAGGTTA; SEQ ID NO: 21: ACAAGGAGCCTGGGGAT; SEQ ID NO: 22: TGCATGGGTCCCAGTGA; SEQ ID NO: 23: TTGACCATTACCACATTGGTGTGC; and SEQ ID NO: 24: ACCATCGACCAGCACATC

When an amplification method includes the use of two primers, a first primer and a second primer, such as in the polymerase chain reaction, the first primer may be selected from the group consisting of SEQUENCE ID NOS: 17, 18, 21, and 23; and the second primer may be selected from the group consisting of SEQUENCE ID NOS: 16, 19, 20, 22, and

- 26 -

24. Any primer pairs which transcribe nucleic acids toward each other and which are specific for sulfonylurea receptor may be used in accordance with the methods of the present invention.

5           Total extraction of RNA is preferably carried out. As used herein, the term "amplification" refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial  
10 concentration, or in an increase in the concentration of a detectable signal. As used herein, the term template-dependent process is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to  
15 nucleic acid synthesis of an RNA or DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, J. D. et al., In: Molecular Biology of the Gene, 4th Ed., W. A. Benjamin, Inc.,  
20 Menlo Park, Calif. (1987)). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by  
25 Cohen et al. (U.S. Pat. No. 4,237,224), Maniatis, T. et al., Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory, 1982.

          A number of template dependent processes are available to amplify the target sequences of interest present  
30 in a sample. One of the best known amplification methods is the polymerase chain reaction (PCR) which is described in detail in U.S. Patents 4,683,195, 4,683,202 and 4,800,159, and in Innis et al., *PCR Protocols*, Academic Press, Inc., San Diego CA, 1990, each of which is incorporated herein by  
35 reference in its entirety. Briefly, in PCR, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An



- 27 -

excess of deoxynucleoside triphosphates are added to a reaction mixture along with a DNA polymerase (e.g., Taq polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction products and the process is repeated. Preferably a reverse transcriptase PCR amplification procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in EPA No. 320,308, incorporated herein by reference in its entirety. In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Application No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA which has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence which can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the

- 28 -

amplification of target molecules that contain nucleotide 5'-[ alpha -thio]triphosphates in one strand of a restriction site (Walker, G. T., et al., *Proc. Natl. Acad. Sci. (U.S.A.)* 1992, 89:392-396, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, i.e. nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and which involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sulfonylurea receptor specific nucleic acids can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-sulfonylurea receptor specific DNA and middle sequence of sulfonylurea receptor specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe identified as distinctive products, generate a signal which is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a sulfonylurea receptor specific expressed nucleic acid.

Still other amplification methods described in GB Application No. 2 202 328, and in PCT Application No. PCT/US89/01025, each of which is incorporated by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR like, template and enzyme dependent synthesis. The primers may be modified by labelling with a capture

- 29 -

moiety (e.g., biotin) and/or a detector moiety (e.g., enzyme). In the latter application, an excess of labelled probes are added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically.

- 5 After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labelled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh D., et  
10 al., *Proc. Natl. Acad. Sci. (U.S.A.)* **1989**, 86:1173, Gingeras T. R., et al., PCT Application WO 88/10315, each of which is incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification  
15 by standard phenol/chloroform extraction, heat denaturation of a clinical sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer which has sulfonylurea receptor  
20 specific sequences. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat denatured again. In either case the single stranded DNA is made fully double stranded by addition of second sulfonylurea receptor specific primer, followed by  
25 polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into double stranded DNA, and transcribed once against with a polymerase such as T7 or SP6. The resulting  
30 products, whether truncated or complete, indicate sulfonylurea receptor specific sequences.

Davey, C., et al., European Patent Application Publication No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process  
35 involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA ("dsDNA") which may be used in accordance with the present invention. The ssRNA

- 30 -

is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

Miller, H. I., et al., PCT application WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; i.e. new templates are not produced from the resultant RNA transcripts. Other amplification methods include "race" disclosed by Frohman, M. A., In: *PCR Protocols: A Guide to Methods and Applications 1990*, Academic Press, N.Y.) and "one-sided PCR" (Ohara, O., et al., *Proc. Natl. Acad. Sci. (U.S.A.)* **1989**, 86:5673-5677), all references herein incorporated by reference in their entirety.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby

- 31 -

amplifying the di-oligonucleotide (Wu, D. Y. et al., *Genomics* 1989, 4:560, incorporated herein by reference in its entirety), may also be used in the amplification step of the present invention.

5           Test fragment and control fragment may be amplified by any amplification methods known to those of skill in the art, including and not limited to the amplification methods set forth above. For purposes of the present invention, amplification of sequences encoding patient and wild type  
10   sulfonyleurea receptor includes amplification of a portion of a sequence such as and not limited to a portion of the sulfonyleurea receptor sequence of SEQ ID NO: 26, such as sequence of a length of about 10 nucleotides to about 1,000 nucleotides, more preferably about 10 nucleotides to about  
15   100 nucleotides, or having at least 10 nucleotides occurring anywhere within the SEQ ID NO: 26, where sequence differences are known to occur within sulfonyleurea receptor test fragments. Thus, for example, a portion of the sequence encoding the second nucleotide binding fragment (NBF-2)  
20   region of sulfonyleurea receptor of a patient sample and a control sample may be amplified to detect sequence differences between these two sequences.

Following amplification of the test fragment and control fragment, comparison between the amplification  
25   products of the test fragment and control fragment is carried out. Sequence differences such as and not limited to nucleic acid transition and restriction digest pattern alterations may be detected by comparison of the test fragment with the control fragment. Nucleic acid transition includes and is  
30   not limited to a G to A transition at nucleic acid position 750 of SEQ ID NO: 26. Another nucleic acid transition involves a G to A transition at nucleic acid position 27 of SEQ ID NO: 29.

These nucleic acid transitions lead to restriction  
35   fragment length polymorphisms as exemplified by the altered results following *MspI* and *NciI* restriction digests set forth above. Accordingly, the restriction fragment length

- 32 -

polymorphisms of test fragments may be compared to the restriction fragments of control fragments.

Alternatively, the presence or absence of the amplification product may be detected. The nucleic acids are  
5 fragmented into varying sizes of discrete fragments. For example, DNA fragments may be separated according to molecular weight by methods such as and not limited to electrophoresis through an agarose gel matrix. The gels are then analyzed by Southern hybridization. Briefly, DNA in the  
10 gel is transferred to a hybridization substrate or matrix such as and not limited to a nitrocellulose sheet and a nylon membrane. A labelled probe encoding a sulfonylurea mutation is applied to the matrix under selected hybridization conditions so as to hybridize with complementary DNA  
15 localized on the matrix. The probe may be of a length capable of forming a stable duplex. The probe may have a size range of about 200 to about 10,000 nucleotides in length, preferably about 500 nucleotides in length, and more preferably about 2,454 nucleotides in length. The preferred  
20 sequence of the probe is set forth in SEQ ID NO: 30. Mismatches which permit substantial similarity to SEQ ID NO: 30, such as and not limited to sequences with similar hydrophobicity and hydrophilicity, will be known to those of skill in the art once armed with the present disclosure.  
25 Various labels for visualization or detection are known to those of skill in the art, such as and not limited to fluorescent staining, ethidium bromide staining for example, avidin/biotin, radioactive labeling such as <sup>32</sup>P labeling, and the like. Preferably, the product, such as the PCR product,  
30 may be run on an agarose gel and visualized using a stain such as ethidium bromide. See Sambrook et al., *supra*. The matrix may then be analyzed by autoradiography to locate particular fragments which hybridize to the probe. Yet another alternative is the sequencing of the test fragment  
35 and the control fragment to identify sequence differences. Methods of nucleic acid sequencing are known to those of skill in the art, including and not limited to the methods of

- 33 -

Maxam and Gilbert, *Proc. Natl. Acad. Sci., USA* 1977, 74, 560-564 and Sanger, *Proc. Natl. Acad. Sci., USA* 1977, 74, 5463-5467.

A diagnostic kit for detecting PHHI comprising in  
5 one or more containers at least one primer which is  
complementary to a sulfonylurea receptor sequence and a means  
for visualizing amplified DNA is also within the scope of the  
present invention. Alternatively, the kit may comprise two  
primers. In either case, the primers may be selected from  
10 the group consisting of SEQ ID NOS: 16-24, for example. The  
diagnostic kit may comprise a pair of primers wherein one  
primer within said pair is complementary to a region of the  
sulfonylurea receptor gene, wherein one of said pair of  
primers is selected from the group consisting of SEQ ID NO:  
15 16-24, a probe specific to the amplified product, and a means  
for visualizing amplified DNA, and optionally including one  
or more size markers, and positive and negative controls.  
The diagnostic kit of the present invention may comprise one  
or more of a fluorescent dye such as ethidium bromide stain,  
20 <sup>32</sup>P, and biotin, as a means for visualizing or detecting  
amplified DNA. Optionally the kit may include one or more  
size markers, positive and negative controls, restriction  
enzymes such as and not limited to *MspI* and/or *NciI*, and/or a  
probe specific to the amplified product.  
25 The following examples are illustrative but are not  
meant to be limiting of the invention.

#### EXAMPLES

##### Purification and Partial Characterization of the 140 kDa Receptor:

30 HIT cell membranes were photolabeled using a  
radioiodinated derivative of the second generation  
hypoglycemic drug, glyburide, according to the methods of  
Nelson, D.A., et al., *JBC*, 1992, 267:14928, Aguilar-Bryan,  
L., et al., *JBC*, 1992, 267:14934, and Aguilar-Bryan, L., et  
35 al., *JBC*, 1990, 265:8218, the disclosures of which are hereby  
incorporated by reference in its entirety.

- 34 -

Glyburide (Kramer et al. *FEBS Lett.* **1988** 229:355-359) and an iodinated derivative of glyburide (Aguilar-Bryan et al. *J. Biol. Chem.* **1990** 265:8218-8224) are known to photolabel a 140 kDa polypeptide. The pharmacological characteristics of the photolabeling, a kD in the low nanomolar range, and appropriate rank order of displacement with other insulin-releasing sulfonylureas, are those expected from studies on glyburide-induced insulin release from islets (Panten et al. *Biochem. Pharm.* **1989** 38:1217-1229) and  $\beta$ -cell lines (Schmid-Antomarchi et al. *J. Biol. Chem.* **1987** 262:15840-15844) and inhibition of  $K_{ATP}$  channel activity. Glyburide was purchased from Sigma (St. Louis, MO) and prepared in stock solutions of 10 mM in dimethyl sulfoxide. Radioligand stocks were prepared by diluting high pressure liquid chromatography-purified 5-[<sup>125</sup>I]iodo-2-hydroxyglyburide in dimethyl sulfoxide. Specific activity (cpm/mol) was measured on radioligand diluted 1/1000 into 10mM Tris, 100 mM NaCl, 2 mM EDTA, pH 7.4, and the absorbance determined at 2.5 nm intervals in a UV-VIS Gilford spectrophotometer. Dimethylsulfoxide was diluted 1/1000 into the same buffer, and the absorbance of the buffer without drug was subtracted at each wavelength to generate the final absorbance profile.

HIT cells, passage 67-73, were seeded in roller bottles at  $50 \times 10^6$  cells/bottle in 100 ml of Dulbecco's modified Eagle's medium plus 10% fetal bovine serum. Cells were fed with 200 ml of medium plus serum 4-5 times over a period of 2 weeks until the cells were confluent. After plating and each feeding, bottles were gassed with 5% CO<sub>2</sub> prior to capping.

The cells in confluent roller bottles were washed with phosphate-buffered saline (0.14 M NaCl, 3 mM KCl, 2 mM KH<sub>2</sub>PO<sub>4</sub>, 1 mM Na<sub>2</sub>HPO<sub>4</sub>, pH 6.8) and then incubated at room temperature with 25 ml of phosphate-buffered saline plus 2 mM EDTA until cells detached from the sides of the bottles. Cells were pelleted at 900 xg for 10 minutes at 4 °C.

All steps were carried out at 0-4 °C. Cell pellets were resuspended in 5 mM Tris, 2 mM EDTA, 0.1 mM PMSF, pH



- 35 -

7.4, using approximately 5 ml of buffer for each roller bottle. Cells were placed on ice for 40 minutes to allow swelling and then homogenized with 10 strokes of a motorized glass-Teflon homogenizer (500 rpm). The homogenate was  
5 centrifuged at 1000 xg for 10 minutes to remove nuclei and cellular debris, and the supernatant transferred to 30 ml of Beckman polycarbonate, screw-cap ultracentrifuge tubes. Supernatants were centrifuged at 100,000 xg for 60 minutes in a Beckman 60 Ti rotor. The pellets were resuspended in  
10 membrane storage buffer (10 mM Tris, 100 mM NaCl, 2 mM EDTA, 20% glycerol, 0.1 mM PMSF, pH 7.4). 200 mg of membrane protein were typically obtained from 20 roller bottles.

Membranes were stored at -80° C at 5 mg/ml protein in 10 mM Tris (pH 7.5), 0.1 M NaCl, 2 mM EDTA, 20% glycerol.  
15 To monitor receptor purification, an aliquot (5-20 ml) of the membranes was incubated with 1 nM [<sup>125</sup>I]-iodo-2-hydroxyglyburide for 15 minutes and the sample photolabeled. Binding of 5-[<sup>125</sup>I]-iodo-2-hydroxyglyburide (5-10 nM) to membranes was done for 30 minutes at 23 °C. Aliquots were  
20 pipetted onto parafilm and irradiated at 23 °C in a UV cross-linker (Fisher Scientific). The energy settings for the UV cross-linker were factory calibrated at 254 nm. For cross-linking at 312 nm, a conversion factor was estimated by determining the time required for the UV cross-linker to  
25 deliver a specific amount of energy with each set of bulbs, and then multiplying by the ratio of these times.

All subsequent steps were performed at room temperature in the presence of 0.1 mM PMSF, 0.1 mM phenanthroline and 0.1 mM iodoacetamide. 20% (w/v) digitonin  
30 was freshly prepared by boiling in deionized water, then added to 200-400 mg thawed labeled membranes to a final concentration of 1%. Membranes were solubilized for 15 minutes then sedimented for 1 hr at 100,000 xg. The supernatant was divided into 4 ml aliquots and each aliquot  
35 was chromatographed over a 1 ml Concanavalin A-Sepharose column equilibrated with 25 mM Tris-HCl, pH 7.5, 0.1 M NaCl, 2 mM EDTA, 1% digitonin. The solution was cycled through the

- 36 -

column twice before washing the column with 8 ml of the equilibration buffer. Retained protein was eluted with 4 ml of the same buffer containing 0.5 M methyl  $\alpha$ -D-mannopyranoside. The eluted protein was stored at  $-80^{\circ}\text{C}$ .

5 Three Con A eluates were combined, then cycled twice over a 1 ml column of reactive green 19-agarose equilibrated with 50 mM HEPES (pH 8.5), 2 mM EDTA, 0.2% digitonin. The column was washed with 8 ml of the equilibration buffer followed by 8 ml of the same buffer containing 0.4 M NaCl. The retained  
10 protein was eluted with 4 ml of the equilibration buffer containing 1.5 M NaCl. The two pooled eluates were diluted 1:1 with the HEPES equilibration buffer without NaCl and cycled twice over a 1 ml phenylboronate-10 agarose column. The column was washed with 8 ml of the HEPES buffer, followed  
15 by 2 ml of 0.1 M Tris-HCl, pH 7.5, 2 mM EDTA, 0.1% digitonin. Protein was eluted with 4 ml of 0.1 M Tris (pH 7.5), 2 mM EDTA, 0.1 % SDS. The protein was concentrated to 0.5 ml using a 100,000 MW cutoff Amicon filter, pretreated with 5% Tween-20, then loaded onto a single 5 cm wide lane of a 5.5%  
20 polyacrylamide SDS gel. After electrophoresis the gel was stained with Coomassie blue, destained, and the receptor band excised with a razor blade. The receptor was electroeluted into a 14,000 MW cutoff dialysis bag and concentrated by Amicon filtration.

25 Table 1 summarizes the yields and fold-purification in the scheme developed for receptor purification. The amount of receptor, yields, and fold-purification reported after each step are based on the radioactivity, determined by  $\gamma$  counting, in the 140 kDa band after electrophoresis  
30 relative to the total protein loaded on a gel lane (as determined using the BioRad protein assay). HIT cell membranes contain approximately 1.6 pmol of receptor per mg of membrane protein as determined by filtration binding (Aguilar-Bryan et al. *J. Biol. Chem.* 1990 265:8218-8224).

- 37 -

TABLE 1  
Purification of the High Affinity 140 kDa Sulfonyleurea Receptor from HIT cells

Step	Total Volume ml	Total Protein mg	Receptor pmol	Receptor pmol/mg	Purification - fold	Yield %
5 Crude Membranes	90	200	320	1.6	1	100
Supernatant	90	150	240	1.6	1	75
ConA-Sepharose	48	10.2	80	7.8	4.9	25
10 Reactive Green 19-agarose	16	1.8	56	31.1	19.5	18
Phenyl boronate agarose	4	0.56	45	80.4	50.4	14
15 SDS-PAGE electroelute	0.2	~0.002	8	4000	2507	2.5

For the autoradiogram depicted in Figure 1, 1-2  $\mu$ g of purified, radiolabeled receptor was made 1% in  $\beta$ -octylglucoside and divided into 6 aliquots. Lane 1 contained receptor kept on ice. The receptor was incubated in the presence (lane 2) and absence (lane 3) of Endo F for 30 min at 37 °C. Aliquots of the samples for lanes 1-3 were further incubated with V8 protease (1  $\mu$ g/10  $\mu$ l) for 30 min at 37 °C, yielding two radiolabeled peptides of 66 and 49 kDa (lanes 4 and 6), both of which are N-glycosylated as indicated by the mobility shift after endo F treatment (lane 5). To obtain N-terminal sequence from the intact receptor, 2  $\mu$ g of protein was separated by electrophoresis on a single, 0.8 cm wide lane of a 5.5% gel. The receptor was transferred to ProBlot (Applied Biosystems) in 10 mM CAPS (pH 11), 10% MeOH, the filter stained for 10-20 seconds with Coomassie blue, destained, the band excised and microsequenced. To prepare receptor fragments for microsequencing, 10  $\mu$ g of purified receptor was cleaved with V8, electrophoresed on a single lane and the fragments from the partial digest transferred to ProBlot. Fragments were prepared and sequenced multiple times as indicated in the figure. Gels used in the

- 38 -

preparation of receptor and fragments for microsequencing were aged overnight, and the top tray buffer contained 0.1 mM thioglycolate.

The purified receptor showed a small apparent  
5 molecular weight decrease ( $\Delta M_r \sim 3000$ ) following treatment with  
Endoglycosidase F/N-glycosidase F (Endo F) and yielded two  
bands following limited cleavage with V8 protease (Figure 1).  
Each of the major labeled proteolytic fragments,  $M_r \sim 69$  and 49  
kDa, shift mobility after digestion with Endo F. Identical  
10 N-terminal sequence, 15-25 residues, were recovered from each  
of the major labeled peptides. No residue was obtained at  
residue 9 when the glycosylated peptides were sequenced; an  
aspartic acid was identified at residue 9 (see Figure 3) in  
the deglycosylated receptor indicating this is an N-  
15 glycosylated asparagine. In addition, N-terminal sequences  
were recovered on two unlabeled V8 peptides and a third minor  
labeled peptide. The results indicate there is an N-linked  
glycosyl group at residue nine in the mature receptor,  
suggesting that the N terminus is extracellular, and that the  
20 sulfonylurea labeling site is within the first 50 kDa of the  
receptor.

Two multiple antipeptide antibodies (MAPs),  
directed against residues 1 through 8 and 10 through 20  
respectively, of Figure 3, both immunoprecipitate  
25 photolabeled 140 kDa receptors from HIT, mouse  $\alpha$ TC-6, and rat  
insulinoma (RIN) cells. MAPs were prepared by synthetic  
protein sequencing (Perkin Elmer-ABI, 430 A Peptide  
Synthesizer, Foster City, CA) to obtain antibodies to M-P-L-  
A-F-C-G-T, residues 1-8 of SEQ ID NOS: 5 and 7. This process  
30 was repeated for residues 10-20 of SEQ ID NOS: 5 and 7, N-H-  
S-A-A-Y-R-V-D-Q-G. A purified sulfonylurea receptor protein  
was immunoprecipitated from HIT cells using the MAPs prepared  
as set forth above.

HIT cell membranes were incubated with 5-[ $^{125}$ I]iodo-  
35 2-hydroxyglyburide, photolabeled, solubilized with 1%  
digitonin, centrifuged at 100,000 xg and the supernatant  
incubated with 1/10 volume of preimmune serum, immune serum,

- 39 -

immune serum + anti-MAP 10-20, or immune serum + irrelevant MAP peptide. 50  $\mu$ l of protein A-Sepharose was added and the mixture was incubated for 2 hours at room temperature, the beads washed with buffer, heated in the presence of pH 9 sample buffer, eletrophoresed on a 6% polyacrylamide SDS gel, and an autoradiogram prepared.

The immunoprecipitation was competed using the immunizing peptide, but not the other MAPS (Figure 2). The amino acid sequence is derived from the photolabeled protein and the N-terminal amino acid sequence is conserved between mouse, rat, and hamster.

#### Isolation and Characterization of cDNA Clones:

Degenerate PCR primers with flanking restriction sites were designed based on the sequence obtained from the labeled peptides. The primers used were as follows:

primer 1 (SEQ ID NO: 10):

5' GAGAGAAGCTT(T/C)TG(T/C)GG(T/C/G/A)GA(A/G)AA(T/C)CA-3'

primer 2 (SEQ ID NO: 11):

5' GAGAGAGAATTCC(T/C)TG(A/G)TC(T/C/G/A)AC(T/C/G/A)C(G/T)(A/G)TA-3'

The bases in parenthesis indicate the degeneracy at that position. The sequence in bold was derived from the peptide sequence obtained from the N-terminus of the sulfonylurea receptor. The remaining 5' sequence was added to facilitate subcloning. Primer 1 has a HindIII site at the 5' end; Primer 2 was engineered with an EcoRI site at the 5' end. These primers were used in a standard PCR reaction with a random primed cDNA library, constructed in  $\lambda$ ZAPII using mouse  $\alpha$ -cell poly A+ mRNA, as template. The following cycle times and temperatures were employed: 94 °C for 10 minutes; 85 °C for 3 minutes, [50 °C for 2 minutes, 72 °C for 5 minutes, 94 °C for 2 minutes,] 50 °C for 2 minutes, 72 °C for 5 minutes.

The bracketed conditions were cycled 30 times. From the N-terminal peptide sequence of the receptor a 47 base pair coding region was expected to be amplified plus the 20 base pairs added to the primers to facilitate cloning

- 40 -

yielding an expected 67 base pair product. The 47 base pair coding region was predicted to have 14 base pairs that were not present in the primers. The PCR product obtained was approximately 67 base pairs and was restricted with EcoRI and HindIII, subcloned into M13 and sequenced. The resulting sequence gave the expected 14 base pairs indicating the sequence was derived from the receptor. The 47 base pair oligonucleotide given below was synthesized based on the consensus sequence derived from nine M13 clones:

5' TTTTGCGGGACGGAGAATCACTCGGCCGCCTACCGCGTCGACCAAGG-3'  
(SEQ ID NO: 12). This oligonucleotide was used to screen the random primed mouse  $\alpha$ TC-cell cDNA library.

A 1.1 kb cDNA was cloned which encoded 28 amino acids obtained from peptide sequencing. This cDNA fragment was used to screen RIN and HIT cell cDNA  $\lambda$  libraries to obtain full sequence.

The nucleotide sequence of a 4635 bp rat receptor cDNA is given in Figure 3. The open reading frame encodes a 1498 amino acid protein with a mass of 167,834 daltons, larger than predicted by SDS polyacrylamide gel electrophoresis. Aguilar-Bryan, L., et al., *JBC*, 1990, 265:8218. The amino acid sequence of the hamster receptor, immediately below the rat sequence in Figure 3, is approximately 98% identical. There is a single insertion of an asparagine at position 742 and a deletion of a threonine at position 831. The first difference between the hamster and rat sequences is in the same relative position, 21 residues C-terminal of the Walker consensus site, as the  $\Delta$ F508 deletion seen in a common cystic fibrosis transconductance regulator (CFTR) mutation (Riordan et al. *Science* 1989 245:1066-1073).

In addition to the insertion and deletion, the first nucleotide binding fold contains approximately a third (10/33) of all differences between the two species.

The mature rat protein, defined by peptide sequencing, begins with a proline following the methionine start site. In the RIN cell receptors the adjacent amino

- 41 -

acid is a methionine. This is the initiating methionine based on the surrounding sequence which is a good fit to the consensus pattern for initiation, GCC(A/G)CCAUG(G) (SEQ ID NO: 13) (Kozak, M. *Cell* **1986** 44:283), including the strongly  
5 conserved A at position -3. However, in the mouse receptor, an additional 35 amino acids is found preceding this proline which cannot eliminate the possibility that some forms of the hamster and rat receptors have similar leader sequences. Confirming the chemical sequence, residue 9 in the mature  
10 proteins is an asparagine within a consensus glycosylation site.

A Blast search of the National Center for Biotechnology Information (NCBI) nucleotide database with the receptor sequence produced matches with several members of  
15 the P-glycoprotein/multidrug resistance protein family. A similar search with the amino acid sequence indicated the sulfonylurea receptor is a member of the ATP-binding cassette superfamily with two putative nucleotide binding domains. The sulfonylurea receptor sequence revealed 29% similarity,  
20 to an ATP-binding cassette superfamily member, termed a multidrug resistance-associated protein (MRP), isolated from a small cell lung carcinoma cell line (H69AR) selected with doxorubicin (Cole et al. *Science* **1992** 258:1650-1654) see Figure 4. A cluster analysis of this molecule, dvhuar in the  
25 Protein Identification Resource (PIR) database, indicates it is related to the leishmania P-glycoprotein-related molecule (Lei/Pgpa), the CFTRs (human (Hum/CFTR), bovine (Bov/CFTR), mouse (Mus/CFTR), and dogfish (Squ/CFTR)) (Cole et al. *Science* **1992** 258:1650-1654). A similar result was obtained  
30 for the sulfonylurea receptor with the additional inclusion of the *Xenopus* CFTR indicating the receptor is a member of this cluster.

The identification of the nucleotide binding domains goes beyond simply having Walker "A" and "B"  
35 consensus sequences. The receptor is similar to the 230-240 amino acid nucleotide binding domain(s) described by (Hyde et al. *Nature* **1990** 346:362-365) and database searches find

- 42 -

similarities to the nucleotide binding fold of ATP-binding proteins. The more conserved of the two receptor nucleotide binding folds, based on similarity with other ATP-binding proteins and the comparison of the rat and hamster sequences, is at the C-terminal end. An alignment of this region with four of the cystic fibrosis transconductance regulator (CFTRs) and multidrug resistance proteins (MRPs) with similarity to the sulfonylurea receptor are shown in Figure 5.

#### 10 RNA Analysis:

Northern blot analysis of poly A+ mRNA isolated from RIN, HIT and  $\alpha$ TC-6 cells, previously shown to have the high affinity receptor by drug binding and photolabeling studies (Aguilar-Bryan et al *J. Cell. Biochem. Suppl.* 1994 18A:133) each have an approximately 5000 nucleotide transcript, see Figure 6. A preliminary tissue distribution study shows the same size transcript is present in mouse brain and heart.

#### Predicted Protein Structure:

Sequence similarities indicate the sulfonylurea receptor has two potential ATP binding folds. The size and additional sequence similarities with P-glycoproteins and CFTRs suggest the receptor has a similar structure. Hydrophobicity (Figure 7) and hydrophobicity versus hydrophobic moment (Eisenberg et al. *J. Mol. Biol.* 1984 179:125) plots were used to generate a model for the receptor (Figure 8). Two constraints were imposed on the model structure: the glycosylation site is on the external face of the membrane and both nucleotide binding domains are on the internal face. The 'classical' ATP-binding cassette superfamily model proposes duplication of a unit consisting of six transmembrane spanning helices followed by a nucleotide binding domain. The sulfonylurea receptor differs from this model and has at least nine potential transmembrane helices before the first nucleotide binding domain but only



- 43 -

four between the two nucleotide binding domains (Figure 8). The multidrug resistance-associated protein (MRP) is predicted to have 8 transmembrane spanning helices (Cole et al. *Science* **1992** 258:1650-1654).

5           Phosphorylation has been implicated in regulation of  $K_{ATP}$  channel activity (Schwanstecher et al. *J. Pharmacol. Exper. Ther.* **1992** 262:495-502) and has been proposed to change the affinity of the sulfonylurea receptor for various ligands. There are 21 potential phosphorylation sites in the  
10 receptor sequence; 3 protein Kinase A (pKA) sites and 18 protein kinase C (pKC) sites. The pKA site at 278 is predicted to be on the external face of the membrane, while those at positions 1363 and 1417 are in the second nucleotide binding fold. Four of the pKC sites (positions 151, 200, 304  
15 and 1213) are predicted to be extracellular or in a membrane spanning helix. Seven of the remaining 14 are in the nucleotide binding folds (NBF); 4 in NBF-1, and 3 in NBF-2. One of the latter sites, Thr 1297 in the Walker A consensus site, is expected to alter nucleotide binding if it is  
20 accessible for phosphorylation.

#### Functional Properties, In Vitro Translations:

mRNA, transcribed by SP6 RNA polymerase from the rat cDNA subcloned into pGEM4, was translated *in vitro*. Approximately 0.5  $\mu$ g of mRNA was heated to 70 °C for 10  
25 minutes, immediately cooled on ice then added to rabbit reticulocyte lysate (Promega, Madison, WI) supplemented with ribonuclease inhibitor, an amino acid mixture, and [<sup>35</sup>S]methionine. The reaction mixture was incubated at 30 °C for 60 minutes then aliquots were subject to electrophoresis  
30 on SDS polyacrylamide gels using standard protocols. The gels were dried and autoradiographed.

The resulting protein was approximately 137 kDa, indicating the receptor behaves anomalously on SDS polyacrylamide gels having a faster than expected mobility,  
35 see Figure 9a. A similar anomalous behavior has been reported for CFTRs (Gregory et al. *Nature* **1990** 347:382-386).

- 44 -

**Anti-Nucleotide Fold Antibodies Immunoprecipitate the  
Photolabeled 140 kDa Receptor:**

Antibodies were produced against two fusion proteins containing the two nucleotide binding folds.

- 5 Fragments of the receptor cDNA were subcloned in frame into pMALc (New England BioLabs, Boston, MA) at the C-terminal end of the DNA encoding the maltose binding protein (MBP). A plasmid expressing the first nucleotide binding fold fused to MBP was constructed by restricting pMALc with StuI and SalI  
10 and restricting the sulfonylurea receptor cDNA with PvuII plus XhoI. A unique 500 base pair fragment was gel purified from the receptor cDNA digest and subcloned into pMALc. The construction was verified by sequencing. The receptor segment expressed is leu708 to leu874. Expression was  
15 obtained in *E. coli* following transformation and induction by isopropylthiogalactoside per the manufacturer's directions. The expressed proteins were found to be in inclusion bodies which were solubilized in SDS and separated on SDS polyacrylamide gels, see Figure 9b. The fusion protein was  
20 electroeluted, concentrated, and used as an immunogen. The solubilized protein in 200 µg amounts, with complete, or incomplete Freund's adjuvant, was injected interdermally into rabbits using a standard 2-3 week regimen of bleeding and boosting.

**25 Injection of *Xenopus* Oocytes with Receptor mRNA:**

- mRNA, approximately 50 ng, transcribed as described above, was injected into *Xenopus* oocytes. The injected oocytes were assayed for K<sup>+</sup> channel activity after 1-5 days using both two-electrode and patch clamp methods. New K<sup>+</sup>  
30 currents in the injected oocytes were not detected. Similarly, co-injection of mRNAs transcribed from cDNAs encoding two small inward rectifiers, ROMK1 (Ho et al. *Nature* 1993 362:31-38) or a brain homolog of IRK1 (Kelly et al. *Biophysical J.* 1994 66(2):A109) failed to confer sulfonylurea  
35 sensitivity on these K<sup>+</sup> channels. The results suggest that the 140 kDa receptor does not have intrinsic K<sup>+</sup> channel

- 45 -

activity, or that *Xenopus* oocytes are not an adequate background for their expression.

#### Transfection Experiments:

The sulfonylurea receptor cDNA has been ligated  
5 into eukaryotic expression vectors containing SV40 virus,  
adenovirus and cytomegalovirus (CMV) promoters. These  
plasmids have been transfected into COS cells which do not  
have the high affinity sulfonylurea receptor as determined by  
filtration binding and photolabeling studies. To date  
10 experiments with the SV40 plasmid have shown that the  
transfected cells produce an mRNA of the appropriate size as  
determined by Northern blots with receptor cDNA. Metabolic  
labeling experiments with the SV40 plasmid where transfected  
and non-transfected cells were labeled with [<sup>35</sup>S] methionine  
15 indicate that the transfected, but not the non-transfected  
cells, synthesize an appropriate sized protein which can be  
immunoprecipitated with the antinucleotide binding fold  
antibodies. The level of receptor synthesized by COS cells  
using this promoter has been low using SEQ ID NOS: 4 and 5.  
20 Expression levels are high using SEQ ID NOS: 32, 33, 35, and  
36 from rat and hamster.

#### Chromosomal localization of the Sulfonylurea Receptor Gene

Chromosomal localization of the Sulfonylurea  
Receptor (*SUR*) gene to normal male human banded chromosomes  
25 was determined by utilization of the fluorescence in situ  
hybridization (FISH) technique by staining with 4,6-  
diamidino-2-phenylindole (DAPI). A metaphase spread showed  
the two chromosome 11 homologues which map the *SUR* cDNA to  
11p15.1. Overlapping human *SUR* cDNA plasmids "mid" and "3",  
30 totaling 3.8 kb, were labeled with biotin-14-dATP (GIBCO) and  
hybridized *in situ* to standard metaphase spreads from normal  
male peripheral blood lymphocytes, according to the methods  
of P. Lichter et al., *Science* 247, 64 (1990), the disclosure  
of which is hereby incorporated by reference in its entirety.  
35 The biotin-labeled DNA was detected using Fluorescein-Avidin

- 46 -

DCS (Vector Laboratories, Burlingame, California). Chromosomes were identified by simultaneous DAPI staining, which produces a Q-banding pattern. Fifteen metaphases were analyzed. Digital images were obtained with a cooled charge-  
5 coupled device camera mounted on a standard epifluorescent microscope (Axioplan; Zeiss, Thronwood, New York). Images were acquired using the software ISee (Inovision Co.) running on a Sun workstation. Fluorescein isothiocyanate and DAPI fluorescence were recorded separately as gray scale images  
10 and then merged using the software package NIH 1.55 (J.W. Ijdo, E.A. Lindsay, R.A. Wells, A. Baldini, *Genomics* **14**, 1019 (1992)). Eighty-five per cent of metaphases analyzed showed specific hybridization signal on both chromatids of the two chromosomes 11 at 11p15.1.

15 Partial cDNA clones, comprising 3.8 kb of coding sequence of the human homologue of *SUR*, were obtained from a human pancreatic cDNA library (provided by Graeme Bell, University of Chicago, and commercial libraries of Clontech, Palo Alto, CA and Invitrogen, San Diego, CA). The library  
20 was produced in lambda gt10 phage (Bell RIN library) and screened with a 2294 bp hamster cDNA probe encoded by SEQ ID NO: 31.

The protocol for making the library is provided by Sambrook et al., *supra*. Poly A+ mRNA was isolated using an  
25 oligo dT column. Poly A+ mRNA was incubated with oligo dT and random hexamers plus reverse transcriptase (such as MMLV RT from Promega, Stratagene or NEBL) and dNTPs to produce single strand cDNA. The single strand cDNA is treated with *E. coli* DNA polymerase, RNaseH and dNTPs, then ligated to  
30 linkers that have *EcoRI* sites to produce double stranded DNA. The final product is restricted with *EcoRI* and ligated, using T4 DNA ligase, into lambda phage DNA that has been similarly restricted and dephosphorylated with alkaline phosphatase to prevent self ligation. The ligated product is packaged into  
35 phage using commercially available packaging extracts.

Screening involved plating and hybridizing at 55°C or 65°C in 5X or 6X SSC (according to the methods of

- 47 -

Sambrook, et al.). 55°C was used for cross species screens and 65°C was employed for the same species. Two washes were carried out at room temperature using 2X SSC, then one at the hybridization temperature of 65°C using 0.1X SSC.

5 Hybridizations and washes were done at reduced stringency (55°C) using methods according to F.M. Ausubel et al., *Current Protocols in Molecular Biology* (Greene Publishing Associates, Inc, New York, NY, 1989), Chap. 6, the disclosures of which are hereby incorporated by reference in  
10 their entirety. Subsequent screening was done at higher stringency (65°C), using a human cDNA of SEQ ID NO: 29 obtained from the first screen as a probe.

Characterization of these cDNA clones by sequence analysis revealed an overall homology of 95% with the rat *SUR*  
15 gene. A specific hybridization signal was detected at the band 11p15.1 in 85% of metaphases on both chromatids of the two chromosomes 11.

#### **Detection of Sulfonylurea Receptor Mutations in PHHI Affected Individuals**

20 Mutational analysis was performed on samples from 16 affected progeny of nine consanguineous matings. In each case, diagnosis of PHHI was based on criteria established by A. Aynsley-Green et al., *supra.*, the disclosure of which is hereby incorporated by reference in its entirety. The  
25 parents in six families were first cousins, in two families second cousins, and in one family more distantly related. Eight families were of Saudi Arabian origin, recruited from the patient population of the Arabian American Oil Company Hospital Medical Services Organization, after institutional  
30 approval was received, and one was of Germanic origin. Family labels follow the form of Thomas et al., *supra.*

Studies indicated that no major insertions or deletions of the *SUR* locus had occurred in three of the families. The first region evaluated, by direct sequence  
35 analysis, was the second nucleotide binding fold (NBF-2) of the human *SUR* homologue (Figure 10). This is the most highly

- 48 -

conserved region of the *SUR* gene, and in other superfamily members it, as well as NBF-1, has functional importance for control of channel activity through interaction with cytosolic nucleotides. S.C. Hyde, *Nature* **346**, 362 (1990) and

5 M.J. Weish, A.E. Smith, *Cell* **73**, 1251 (1993).

To obtain this genomic structure, a normal human lymphocyte genomic bacteriophage library (provided by Mary Beth Humphrey, Baylor College of Medicine) was screened, using standard methods according to F.M. Ausubel et al.,  
10 *supra.*, with a human partial *SUR* cDNA probe of SEQ ID NO: 29 (cDNA probe, "3prime").

The human genomic library was made in lambda FIX using materials supplied by Stratagene, Inc. Briefly, genomic DNA was partially digested with *Sau3A*, the fragments  
15 were precipitated with ethanol, resuspended with pre-cut lambda FIX DNA which has compatible ends, ligated with T4 DNA ligase and packaged and screened.

Hybridizations and washes were done at reduced stringency (55°C) using methods according to F.M. Ausubel et  
20 al., *Current Protocols in Molecular Biology* (Greene Publishing Associates, Inc, New York, NY, 1989), Chap. 6, the disclosures of which are hereby incorporated by reference in their entirety. The library was screened with a 1.2 kb hamster cDNA probe of SEQ ID NO: 30, which spans the *SUR* NBF2  
25 sequence. Subsequent screening was done at higher stringency (65°C), using a human cDNA of SEQ ID NO: 29 obtained from the first screen as a probe. Screening involved plating and hybridizing at 55°C or 65°C in 5X or 6X SSC (according to the methods of Sambrook, et al.). 55°C was used for cross  
30 species screens and 65°C was employed for the same species. Two washes were carried out at room temperature using 2X SSC, then one at the hybridization temperature of 65°C using 0.1X SSC.

Inserts in the bacteriophage clone  $\lambda$ G4 were  
35 subcloned into pBluescript 11 (Stratagene, La Jolla, California). Plasmids were purified using standard cesium chloride purified methods, restricted using the appropriate

- 49 -

desired enzyme(s). The fragments were purified by electrophoresis on low melt agarose and cut out of the gel. A 1-to-5 microliter aliquot of the desired fragment and 1 microgram of the appropriately restricted plasmid carrying a selectable ampicillin resistance marker (such as pBluescript from Stratagene, Inc.) were melted at 65°C, mixed and diluted to 20 microliters with a buffer containing T4 DNA ligase and ATP, then incubated for 4-18 hours before transforming into *E. coli* and selecting on ampicillin plates.

Exon-intron boundaries were defined by comparing the nucleotide sequences of the human *SUR* gene and cDNA, which were obtained using the dideoxy chain termination method (Sequenase; U.S. Biochemicals, Cleveland, Ohio).

Because of the consanguineous matings and autosomal recessive inheritance pattern of this disorder, affected individuals are expected to be homozygous by descent at the disease gene locus. E.S. Lander and D. Botstein, *Science* **236**, 1567 (1987), the disclosure of which is hereby incorporated by reference in its entirety. Direct sequencing of a pancreatic cDNA product, isolated from an affected child of Family 6, revealed a 109 bp deletion within the NBF-2 region which corresponded to skipping of an exon resulting in a cDNA product of about 2190 bp in length using primers of SEQ ID NOS: 20 and 21 as compared to mRNA of about 2080 bp in length. The effects of this skipping event are severe and include production of a frameshift, premature truncation of the protein due to inclusion of a stop 24 codons later, and disruption of the NBF-2 (Figure 11A and 11B). The splice sites of the skipped exon were evaluated at the genomic DNA level and a homozygous G to A point mutation, located within the 5' splice site at the last base of the skipped exon, was found (Figure 11C). A recognition site for the restriction endonuclease *MspI* is destroyed by this base change, providing a means to confirm and test for the presence of the mutation. mRNA was directly isolated using Oligotex (Qiagen Inc., Studio City, California) from a fresh-frozen pancreatic tissue sample and reverse transcribed (RT), using random

- 50 -

primers (Invitrogen, San Diego, California), with Superscript 11 (GIBCO-BRL) into cDNA. For cloning of the NBF-2 region, an initial PCR amplification with SEQ ID NOS: 23 (primer 22 (located 5' of 17)) and 19 (primer 29) was followed by a  
5 second amplification of a portion of the reaction with SEQ ID NOS: 16 (primer 17) and 19 (primer 29) using conditions described by P.M. Thomas, G.J. Cote, D.M. Hallman, P.M. Mathew, *Am. J. Hum. Genet.*, in press, *supra*.

PCR products were amplified using hybridization at  
10 60°C for 1 minute, elongated at 72°C for 1 minute and denatured at 93°C for 1 minute for thirty cycles. Hybridization may be carried out at temperatures of between about 55°C to about 65°C. The amplified product was cloned into pCR 11™ vector (Invitrogen, San Diego, CA) and  
15 sequenced, as above. pCR 11™ vector is set forth in Figure 13. The vector was purified using a standard cesium chloride method, restricted using the appropriate desired enzyme(s), the fragments were purified by electrophoresis on low melt agarose and cut out of the gel. A 1-to-5 microliter aliquot  
20 of the desired fragment and 1 microgram of the appropriately restricted plasmid carrying a selectable ampicillin resistance marker (such as pBluescript from Stratagene, Inc.) were melted at 65°C, mixed and diluted to 20 microliters with a buffer containing T4 DNA ligase and ATP, then incubated for  
25 4-18 hours before transforming into *E. coli* and selecting on ampicillin plates. For detection of the mutation in genomic fragments, 100 ng of genomic DNA was amplified using SEQ ID NOS: 18 and 20, primers 28 and 29B, as above except in the presence of PCR buffer N (Invitrogen, San Diego, CA), and  
30 either directly PCR sequenced according to the methods of S. Khorana, R.F. Gagel, G.J. Cote, *Nucleic Acids Res.* 22, 3425 (1994), the disclosure of which is hereby incorporated herein by reference in its entirety, or cut with 5 U of *MspI* (GIBCO-BRL) at 37°C for 2 hours and run on a 10% polyacrylamide gel.  
35 Visualization of products was by silver staining. Both affected children of Family 6 were homozygous, while the parents and two unaffected siblings were found to be



- 51 -

heterozygous, for the mutation (Figure 11D). Preliminary semiquantitative analysis revealed markedly decreased expression of the mutant SUR message upon comparison of patient and age-matched normal control pancreatic samples, suggesting instability of the mutant message.

Thirteen additional affected children, from six families of Saudi Arabian origin and one family of German origin, were found to be homozygous for this mutation, as demonstrated by loss of the *MspI* restriction enzyme recognition site. In all families, homozygous loss of the *MspI* site cosegregated with disease phenotype, and in Families 1-3 and 5 genotype analysis for this mutation agreed with previously reported haplotype data, P.M. Thomas, G.J. Cote, D.M. Hallman, P.M. Mathew, *Am. J. Hum. Genet.*, in press, *supra*. Direct sequencing of PCR-amplified genomic DNA from a representative affected member of each family determined that all exhibited the homozygous G to A mutation.

Genomic DNA from affected and normal individuals was PCR-amplified using the SEQ ID NOS: 16 and 19 and cloned into pRSVhMT2A. Plasmids were purified using standard cesium chloride methods, restricted using the appropriate desired enzyme(s). The fragments were purified by electrophoresis on low melt agarose and cut out of the gel. A 1-to-5 microliter aliquot of the desired fragment and 1 microgram of the appropriately restricted plasmid carrying a selectable ampicillin resistance marker (such as pBluescript from Stratagene, Inc.) were melted at 65°C, mixed and diluted to 20 microliters with a buffer containing T4 DNA ligase and ATP, then incubated for 4-18 hours before transforming into *E. coli* and selecting on ampicillin plates.

Constructs were transfected into the human glioblastoma cell line SNB 19 using Lipofectamine™ (Gibco-BRL, Gaithersburg, MD). RT-PCR analysis was performed, with SEQ ID NOS:18 (primer 16) and 23 (primer DS8), as described by H. Lou, G.J. Cote, R.F. Gagel, *Mol. Endo.* 8, 1618 (1994), the disclosure of which is incorporated herein by reference in its entirety. The plasmids and their cDNA products were

- 52 -

sequenced with SEQ ID NO: 17 (primer 34a1). Genomic DNA fragments were PCR-amplified with SEQ ID NOS: 17 and 18 (primers 34a1 and 16) and digested with *NciI*, as in Figure 11.

- 5                   Family 4 demonstrated a unique mutation in the 3' splice site sequence preceding the start of the NBF-2 (Figure 12A). This G to A mutation destroys an *NciI* restriction endonuclease site and homozygous loss of this site cosegregated with disease phenotype within the family.
- 10 Again, genotype analysis of the members of this family supported previously reported haplotype data, P.M. Thomas, G.J. Cote, D.M. Hallman, P.M. Mathew, *Am. J. Hum. Genet.*, in press, *supra.*; both parents are heterozygotes for the mutation and the unaffected sibling is homozygous for the
- 15 wild type allele (Figure 12B). Since a pancreatic tissue sample from an affected individual in Family 4 was unavailable and we were unable to recover the *SUR* message from transformed lymphocytes, a chimeric construct was created to examine the effects of this mutation on the RNA
- 20 splicing pathways according to the methods of R. Takahashi, et al., *Nature Genet.* 7, 79 (1994); I. Satokata, et al., *Proc. Natl. Acad. Sci.* 87 9908 (1990); H. Lou, G.J. Cote, R.F. Gagel, *Mol. Endo.* 8, 1618 (1994), the disclosure of each hereby incorporated by reference in its entirety. With the
- 25 construct containing the mutation, no wild type splicing pattern occurred. Instead, use of three cryptic 3' splice sites was demonstrated resulting in a 7 bp addition, a 20 bp deletion, and a 30 bp deletion in the exon (Figure 12D). A similar intronic 3' splice acceptor mutation, described in
- 30 the disorder 21-hydroxylase deficiency, also resulted in lack of the wild type splicing pattern, produced several cryptic splice products, and abolished normal protein activity. Y. Higashi, et al., *Proc. Natl. Acad. Sci., USA* 85, 7486 (1988), the disclosure of which is incorporated herein by reference
- 35 in its entirety.

All PCR products prepared from genomic DNA of 100 normal, unrelated individuals showed normal *MspI* and *NciI*

- 53 -

restriction patterns, indicating that neither mutation is a common polymorphism. The data presented provides evidence that mutations in the *SUR* gene cause familial persistent hyperinsulinemic hypoglycemia of infancy.

- 5           Various modifications of the invention in addition to those shown and described herein will be apparent to those skilled in the art from the foregoing description. Such modifications are also intended to fall within the scope of the appended claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT:
- (ii) TITLE OF INVENTION: Sequence Encoding Mammalian Sulfonylurea Receptor and Method of Detecting Persistent Hyperinsulinemic Hypoglycemia of Infancy
- (iii) NUMBER OF SEQUENCES: 37
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
  - (B) STREET: One Liberty Place 46th. Floor
  - (C) CITY: Philadelphia
  - (D) STATE: PA
  - (E) COUNTRY: USA
  - (F) ZIP: 19103
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: PCT/US95/04463
  - (B) FILING DATE: 12-APR-1995
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: Beardell, Lori Y.
  - (B) REGISTRATION NUMBER: 34,293
  - (C) REFERENCE/DOCKET NUMBER: BYLR-0004
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 215-568-3100
  - (B) TELEFAX: 215-568-3439

## (2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4599 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGCCTTTGG CTTTCTGCGG CACCGAGAAC CACTCGGCCG CCTACCGGGT GGACCAAGGC	60
GTCTCTCAACA ACGGCTGCTT CGTGGACGCG CTCAATGTGG TGCCACATGT CTTTCTGCTC	120

TTCATCACCT	TCCCCATCCT	CTTCATCGGA	TGGGGCAGCC	AGAGCTCCAA	GGTGACATT	180
CACCACAGCA	CCTGGCTCCA	TTTCCCGGGG	CACAACCTGC	GCTGGATCCT	GACCTTCATA	240
CTGCTCTTCG	TCCTCGTGTG	TGAGATCGCT	GAGGGTATCC	TGTCTGACGG	GGTGACAGAA	300
TCCCCGCCACC	TCCACTTATA	CATGCCAGCT	GGGATGGCAT	TCATGGCTGC	CATCACCTCT	360
GTGGTCTACT	ACCATAACAT	TGAGACCTCT	AACTTTCCCA	AGCTGCTGAT	TGCTCTGCTC	420
ATCTACTGGA	CCCTGGCCTT	CATCACGAAG	ACCATCAAGT	TCGTCAAGTT	CTACGACCAC	480
GCCATTGGCT	TCTCTCAGCT	GCGCTTCTGC	CTCACGGGGC	TTCTGGTGAT	CCTCTACGGG	540
ATGCTGCTGC	TTGTGGAGGT	CAATGTCATC	CGGGTGAGGA	GATACNTCTT	CTTCAAGACA	600
CCAAGGAAG	TAAAGCCCCC	CGAGGACCTA	CAGGACCTGG	GTGTGCGCTT	TCTGCAGCCC	660
TTCGTTAACC	TGCTATCAAA	GGGGACCTAC	TGGTGGATGA	ATGCCTTCAT	CAAGACTGCT	720
CACAAGAAGC	CCATCGACCT	GCGGGCCATC	GNGAAGCTGC	CCATTGCCAT	GAGAGCCCTC	780
ACCAACTACC	AGCGACTCTG	CNTGGCCTTC	GATGCCCAGG	CGCGGAAGGA	CACACAGAGC	840
CNGCAGGGTG	CCCGGGCCAT	CTGGAGGGCT	CTCTGTCATG	CCTTTGGGAG	ACGGCTGGTC	900
CTCAGCAGCA	CATTCCGTAT	CCTGGCCGAC	CTCCTGGGCT	TTGCTGGGCC	ACTCTGCATC	960
TTCGGGATCG	TGGACCACCT	CGGGAAGGAG	AACCACGTCT	TCCAGCCCCA	GACACAGTTT	1020
CTTGAGATTT	ACTTTGTCTC	ATCCCAAGAG	TTCCTCGGCA	ATGCCTATGT	CTTGCTGTT	1080
CTTCTGTTCC	TGCCCCTCCT	GCTGCAAAGG	ACCTTTCTAC	AAGCCTCGTA	CTACGTTGCC	1140
ATTGAAACTG	GGATCAACCT	GAGAGGAGCA	ATCCAGACCA	AGATTTACAA	TAAGATCATG	1200
CACNTGTCTA	CTTCCAACCT	GTCCATGGGG	GAAATGACTG	CTGGGCAGAT	CTGCAACCTG	1260
GTGGCCATCG	ACACCAACCA	GCTCATGTGG	TTTTTCTTCT	TATGCCCCAA	CCTCTGGNCT	1320
ATGCCGGTAC	AGATCATTGT	GGGCGTGATC	CTCCTCTACT	ACATCCTTGG	GGTCAGCGCC	1380
TTGATTGGAG	CGGCTGTCTAT	CATTCTGCTG	GCTCCTGTAC	AGTACTTTGT	GGCCACCAAG	1440
CTGTCCCAGG	CACAGCGGAC	GACCCTGGAA	TATTCCAATG	AGAGGCTGAA	GCAGACCAAT	1500
GAGATGCTCC	GGGGCATNAA	GTTGCTCAAG	CTCTATGCGT	GGGAGAACAT	CTTCTGCTCC	1560
AGGGTGAGAG	AGACACGCAG	GAAGGAAATG	ACCAGCCTCA	GGGCCTTCGC	TGTCTACACC	1620
TCCATCTCCA	TCTTCATGAA	CACAGCTATC	CCCATCGCTG	CTGTCCTCAT	CACCTTCGTG	1680
GGCCACGTCA	GCTTCTTCAA	AGAGTCGGAC	NTCTCGCCCT	CGGTGGCCTT	TGCTCTCTC	1740
TCTCTCTTCC	ACATCCTGGT	CACACCGCTG	TTCCTGCTGT	CTAGTGTGGT	TCGGTCCACT	1800
GTCAAGGCCC	TGGTGAGCGT	GCAAAAGCTG	AGTGAGTTCC	TGTCCAGTGC	AGAGATCCGT	1860
GAGGAACAGT	GTGCCCCCCG	AGAGCCCGCA	CCCCAAGGCC	AAGCGGGCAA	GTACCAGGCG	1920
GTGCCCCCTCA	AGTTCGTAAA	CCGCAAGCGC	CCAGCCCGAG	AAGAAGTCCG	GGACCTCTTG	1980
GGCCCACTGC	AGAGGCTGNC	TCCCAGCANG	GATGGAGACG	CTCACAACCT	CTGTGTCCAG	2040
ATCATCGGAG	GCTTCTTCAC	CTGGACCCCT	GATGGAATCC	CCACCCTGTC	CAACATCACC	2100
ATCCGTATCC	CCCAGGTCA	GCTGACCATG	ATCGTGGGGC	AGGTGGGCTG	TGGCAAGTCC	2160

SUBSTITUTE SHEET (RULE 26)

TEGCTCCTTC	TGGCCACCCT	GGGGGAGATG	CAGAAGGTCT	CTGGAGCTGT	CTTCTGGAAC	2220
AGCCTTCCAG	ACAGCGAGGG	GAGANGACCC	CAGCAACCCA	GAGCGGGAGA	CAGCNGCNGN	2280
NTCGGATNCC	AGGAGCAGAG	GGCCNGTGGC	TACGCATCTC	AGAAACCATG	GCTGCTAAAT	2340
GCCACTGTGG	AGGAGAACAT	CACCTTCGAG	AGTCCCTTCA	ATNNGCAACG	GTACAAGATG	2400
GTCATCGAAG	CCTGCTCCCT	GCAGCCAGAC	ATAGACATCC	TGCCCCATGG	AGACCAGACT	2460
CAGATTGGGG	AACGAGGCAT	CAACTTGAGT	ACTGGTGGTC	AGCGTCCAGA	TCAGTGTNGA	2520
CCCGAGCCCT	CTACCAGCAN	ACCAATGNNT	GTCTTTTGG	ATGACCCTTT	CTCGGCTCTG	2580
GATGTCCATC	TGAGTGACCA	CCTAATGCAG	GCTGGCATCC	TCGAGCTGCT	CCGGGATGAC	2640
AAGAGGACAG	TGGTCTTGGT	GACCCACAAG	CTACAGTACC	TGCCTCATGC	TGACTGGATC	2700
ATTGCTATGA	AGGATGGCAC	CATTGAGAGG	GAGGGGACAC	TCAAGGACTT	CCAGAGGTCT	2760
GAGTGCCAGC	TCITTGAGCA	TTGGAAGACC	CTCATGAACC	GGCAGGACCA	AGAGCTGGAG	2820
AAGGAGACAG	TCATGGAGAG	AAAAGCCNCN	GAGCCATCTC	AGGGCCTGCC	CCGTGCCATG	2880
TCCTCAAGAG	ATGGCCTTCT	GCTGGATGAG	GANGAGGAGG	AAGAGGAGGC	AGCCGAGAGC	2940
GAGGAAGATG	ACAACCTTATC	CTCTGTGCTG	CATCAGCGAG	CCAAGATCCC	ATGGCGAGCC	3000
TGCACCAAGT	ATTTGTCTCT	TGCTGGCATC	CTGCTCCTGT	CCCTGCTTGT	CTTCTCCCAG	3060
CTGCTCAAGC	ACATGGTCTT	GGTGGCCATT	GACTACTGGC	TGGCCAAGTG	GACGGACAGT	3120
GCCCTGTGTC	TGAGCCCCGC	CGCCAGGAAC	TGCTCCCTCA	GCCAGGAATG	TGNCCTGGAC	3180
CAATCTGTCT	ATGCCATGGT	ATTACACNTG	CTCTGCAGCC	TGGGTATCGN	GCTGTGCCTT	3240
GTCACCTCTG	TCACTGTGGA	GTGGACGGGA	CTGAAGGTGG	CCAAGAGGCT	GCATCGCAGC	3300
CTGCTCAACC	GTATCATCCT	GGCTCCCATG	AGGTTCTTTG	AGACCACGCC	CCTGGGGAGT	3360
ATCCTGAACA	GATTTTCATC	TGACTGTAAC	ACCATTGACC	AGCATATCCC	GTCCACGCTG	3420
GAGTGCCTGA	GCAGATCCAC	CTTACTCTGT	GTCTCCGCCC	TGNCCTGTCAT	CTCCTACGTC	3480
ACGCCTGTGT	TCCTAGTGGC	CCTCTTACCC	CTCGCCGTCG	TGTGCTACTT	CATCCAGAAG	3540
TACTTCCGAG	TGGCGTCCAG	GGACCTGCAG	CAGCTGGACG	ACACAACACA	GCTCCCTCTG	3600
NTCTCACACT	TTGCTGAAAC	TGTGGAAGGA	CTCACCACCA	TCCGTGCCTT	CAGGTACGAG	3660
GCCCCGTTCC	AGCAGAAGCT	CCTAGAGTAC	ACCGACTCCA	ACAACATTGC	CTCTCTCTTC	3720
CTCACAGCAG	CCAACAGGTG	GCTGGAAGTC	NGCATGGAGT	ACATCGGAGC	ATGCGTGGTA	3780
CTCATCGCCG	CTGCCACCTC	CATCTCCAAC	TCCCTACACA	GGGAGCTCTC	AGCCGGCCTA	3840
GTAGGCCTGG	GCCTCACCTA	TGCCTTGATG	ATTGGGATCT	GCGGCCGCAC	AGGCAGTGGA	3900
AAATCCTCCT	TCTCTCTGCG	CTTTTCCGA	ATGGTGGATA	TGTTTGAAGG	GCGTATCATC	3960
ATCGATGGCA	TTGACATCGC	CAAGCTGCCG	CTGCACACGC	TCCGCTCACG	CCTGTCTATC	4020
ATCCTACAGG	ACCCTGTCTT	CTTCAGTGGT	ACCATCAGAT	TCAACCTGGA	CCCAGAGAAG	4080
AAATGCTCAG	ACAGCACGCT	GTGGGAGGCT	CTGGAGATCG	CTCAGCTGAA	GCTGGTGGTG	4140
AAGGCCCTGC	CAGGAGGCCT	GGATGCCATC	ATCACGGAAG	GAGGGGAGAA	TTTTAGCCAG	4200

SUBSTITUTE SHEET (RULE 26)

57

GGCCAGAGGC AGCTGTTCTG CCTGGCCCGG GCCTTTGTGA GGAAGACCAG CATCTTCATC 4260  
ATGGATGAAG CAACTGCCTC CATCGACATG GCTACGGAAA ATATCCTCCA GAAGGTGGTG 4320  
ATGACAGCCT TCGCAGACCG CACCGTGGTC ACCATCGCGC ACCGCGTGCA CACCATCCTG 4380  
AGTGCAGACC TAGTGATGGT CCTGAAGAGG GGC GCGATCC TGGAGTTCGA CAAGCCGGAA 4440  
ANGCTTCTCA GCCAGAAGGA CAGCGTCTTT GCCTCCTTTG TCCGCGCGGA CAAATGACCA 4500  
GCCAGCGCCA AAGTGCCACC CCACACCTCA CCTGCTTGCC ATGGATTCT TACTGTAAAT 4560  
CACTTGTAAT TAAAGAACT AATTCTTTCG TAAAAAAA 4599

## (3) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 36 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GCGCGGAGCC GGAGCCGAGC CCGTGC GCGC GCCACC 36

## (4) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 24 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AGCCGAGCCC GTGCGCGCGC CGCC 24

## (5) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4635 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

SUBSTITUTE SHEET (RULE 26)

(v) FRAGMENT TYPE: N-terminal

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 37..4533

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GCGCGGAGCC GGAGCCGAGC CCGTGC GCGC GCCACC ATG CCT TTG GCC TTC TGC	54
Met Pro Leu Ala Phe Cys	
1 5	
GGC ACC GAG AAC CAC TCG GCC GCC TAC CGG GTG GAC CAA GGC GTC CTC	102
Gly Thr Glu Asn His Ser Ala Ala Tyr Arg Val Asp Gln Gly Val Leu	
10 15 20	
AAC AAC GGC TGC TTC GTG GAC GCG CTC AAT GTG GTG CCA CAT GTC TTT	150
Asn Asn Gly Cys Phe Val Asp Ala Leu Asn Val Val Pro His Val Phe	
25 30 35	
CTG CTC TTC ATC ACC TTC CCC ATC CTC TTC ATC GGA TGG GGC AGC CAG	198
Leu Leu Phe Ile Thr Phe Ile Leu Phe Ile Gly Trp Gly Ser Gln	
40 45 50	
AGC TCC AAG GTG CAC ATT CAC CAC AGC ACC TGG CTC CAT TTC CCG GGG	246
Ser Ser Lys Val His Ile His His Ser Thr Trp Leu His Phe Pro Gly	
55 60 65 70	
CAC AAC CTG CGC TGG ATC CTG ACC TTC ATA CTG CTC TTC GTC CTC GTG	294
His Asn Leu Arg Trp Ile Leu Thr Phe Ile Leu Leu Phe Val Leu Val	
75 80 85	
TGT GAG ATC GCT GAG GGT ATC CTG TCT GAC GGG GTG ACA GAA TCC CGC	342
Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp Gly Val Thr Glu Ser Arg	
90 95 100	
CAC CTC CAC TTA TAC ATG CCA GCT GGG ATG GCA TTC ATG GCT GCC ATC	390
His Leu His Leu Tyr Met Pro Ala Gly Met Ala Phe Met Ala Ala Ile	
105 110 115	
ACC TCT GTG GTC TAC TAC CAT AAC ATT GAG ACC TCT AAC TTT CCC AAG	438
Thr Ser Val Val Tyr Tyr His Asn Ile Glu Thr Ser Asn Phe Pro Lys	
120 125 130	
CTG CTG ATT GCT CTG CTC ATC TAC TGG ACC CTG GCC TTC ATC ACG AAG	486
Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr Leu Ala Phe Ile Thr Lys	
135 140 145 150	
ACC ATC AAG TTC GTC AAG TTC TAC GAC CAC GCC ATT GGC TTC TCT CAG	534
Thr Ile Lys Phe Val Lys Phe Tyr Asp His Ala Ile Gly Phe Ser Gln	
155 160 165	
CTG CGC TTC TGC CTC ACG GGG CTT CTG GTG ATC CTC TAC GGG ATG CTG	582
Leu Arg Phe Cys Leu Thr Gly Leu Leu Val Ile Leu Tyr Gly Met Leu	
170 175 180	
CTG CTT GTG GAG GTC AAT GTC ATC CGG GTG AGG AGA TAC GTC TTC TTC	630
Leu Leu Val Glu Val Asn Val Ile Arg Val Arg Arg Tyr Val Phe Phe	
185 190 195	
AAG ACA CCA AGG GAA GTA AAG CCC CCC GAG GAC CTA CAG GAC CTG GGT	678
Lys Thr Pro Arg Glu Val Lys Pro Pro Glu Asp Leu Gln Asp Leu Gly	
200 205 210	
GTG CGC TTT CTG CAG CCC TTC GTT AAC CTG CTA TCA AAG GGG ACC TAC	726

SUBSTITUTE SHEET (RULE 26)



59

Val 215	Arg	Phe	Leu	Gln	Pro 220	Phe	Val	Asn	Leu	Leu 225	Ser	Lys	Gly	Thr	Tyr 230	
TGG	TGG	ATG	AAT	GCC	TTC	ATC	AAG	ACT	GCT	CAC	AAG	AAG	CCC	ATC	GAC	774
Trp	Trp	Met	Asn	Ala 235	Phe	Ile	Lys	Thr	Ala 240	His	Lys	Lys	Pro	Ile 245	Asp	
CTG	CGG	GCC	ATC	GGG	AAG	CTG	CCC	ATT	GCC	ATG	AGA	GCC	CTC	ACC	AAC	822
Leu	Arg	Ala	Ile 250	Gly	Lys	Leu	Pro	Ile 255	Ala	Met	Arg	Ala	Leu 260	Thr	Asn	
TAC	CAG	CGA	CTC	TGC	TTG	GCC	TTC	GAT	GCC	CAG	GCG	CGG	AAG	GAC	ACA	870
Tyr	Gln	Arg	Leu 265	Cys	Leu	Ala	Phe	Asp	Ala	Gln	Ala	Arg	Lys 275	Asp	Thr	
CAG	AGC	CAG	CAG	GGT	GCC	CGG	GCC	ATC	TGG	AGG	GCT	CTC	TGT	CAT	GCC	918
Gln	Ser	Gln	Gln	Gly	Ala	Arg	Ala	Ile	Trp	Arg	Ala	Leu	Cys	His	Ala	
280							285				290					
TTT	GGG	AGA	CGG	CTG	GTC	CTC	AGC	AGC	ACA	TTC	CGT	ATC	CTG	GCC	GAC	966
Phe	Gly	Arg	Arg	Leu	Val	Leu	Ser	Ser	Thr	Phe	Arg	Ile	Leu	Ala	Asp	
295					300					305					310	
CTC	CTG	GGC	TTT	GCT	GGG	CCA	CTC	TGC	ATC	TTC	GGG	ATC	GTG	GAC	CAC	1014
Leu	Leu	Gly	Phe	Ala 315	Gly	Pro	Leu	Cys	Ile 320	Phe	Gly	Ile	Val	Asp 325	His	
CTC	GGG	AAG	GAG	AAC	CAC	GTC	TTC	CAG	CCC	AAG	ACA	CAG	TTT	CTT	GGA	1062
Leu	Gly	Lys	Glu	Asn	His	Val	Phe	Gln	Pro	Lys	Thr	Gln	Phe	Leu	Gly	
			330					335					340			
GTT	TAC	TTT	GTC	TCA	TCC	CAA	GAG	TTC	CTC	GGC	AAT	GCC	TAT	GTC	TTG	1110
Val	Tyr	Phe	Val	Ser	Ser	Gln	Glu	Phe	Leu	Gly	Asn	Ala	Tyr	Val	Leu	
		345					350				355					
GCT	GTT	CTT	CTG	TTC	CTT	GCC	CTC	CTG	CTG	CAA	AGG	ACC	TTT	CTA	CAA	1158
Ala	Val	Leu	Leu	Phe	Leu	Ala	Leu	Leu	Leu	Gln	Arg	Thr	Phe	Leu	Gln	
360						365					370					
GCC	TCG	TAC	TAC	GTT	GCC	ATT	GAA	ACT	GGG	ATC	AAC	CTG	AGA	GGA	GCA	1206
Ala	Ser	Tyr	Tyr	Val	Ala	Ile	Glu	Thr	Gly	Ile	Asn	Leu	Arg	Gly	Ala	
375					380					385					390	
ATC	CAG	ACC	AAG	ATT	TAC	AAT	AAG	ATC	ATG	CAC	TTG	TCT	ACT	TCC	AAC	1254
Ile	Gln	Thr	Lys	Ile	Tyr	Asn	Lys	Ile	Met	His	Leu	Ser	Thr	Ser	Asn	
				395					400					405		
CTG	TCC	ATG	GGG	GAA	ATG	ACT	GCT	GGG	CAG	ATC	TGC	AAC	CTG	GTG	GCC	1302
Leu	Ser	Met	Gly	Glu	Met	Thr	Ala	Gly	Gln	Ile	Cys	Asn	Leu	Val	Ala	
			410					415					420			
ATC	GAC	ACC	AAC	CAG	CTC	ATG	TGG	TTT	TTC	TTC	TTA	TGC	CCA	AAC	CTC	1350
Ile	Asp	Thr	Asn	Gln	Leu	Met	Trp	Phe	Phe	Phe	Leu	Cys	Pro	Asn	Leu	
			425				430					435				
TGG	GCT	ATG	CCG	GTA	CAG	ATC	ATT	GTG	GGC	GTG	ATC	CTC	CTC	TAC	TAC	1398
Trp	Ala	Met	Pro	Val	Gln	Ile	Ile	Val	Gly	Val	Ile	Leu	Leu	Tyr	Tyr	
440						445					450					
ATC	CTT	GGG	GTC	AGC	GCC	TTG	ATT	GGA	GCG	GCT	GTC	ATC	ATT	CTG	CTG	1446
Ile	Leu	Gly	Val	Ser	Ala	Leu	Ile	Gly	Ala	Ala	Val	Ile	Ile	Leu	Leu	
455					460					465					470	
GCT	CCT	GTA	CAG	TAC	TTT	GTG	GCC	ACC	AAG	CTG	TCC	CAG	GCA	CAG	CGG	1494
Ala	Pro	Val	Gln	Tyr	Phe	Val	Ala	Thr	Lys	Leu	Ser	Gln	Ala	Gln	Arg	
				475					480					485		

SUBSTITUTE SHEET (RULE 26)

60

ACG Thr	ACC Thr	CTG Leu	GAA Glu	TAT Tyr	TCC Ser	AAT Asn	GAG Glu	AGG Arg	CTG Leu	AAG Lys	CAG Gln	ACC Thr	AAT Asn	GAG Glu	ATG Met	1542
		490					495						500			
CTC Leu	CGG Arg	GGC Gly	ATC Ile	AAG Lys	TTG Leu	CTC Leu	AAG Lys	CTC Leu	TAT Tyr	GCG Ala	TGG Trp	GAG Glu	AAC Asn	ATC Ile	TTC Phe	1590
		505					510					515				
TGC Cys	TCC Ser	AGG Arg	GTG Val	GAG Glu	AAG Lys	ACA Thr	CGC Arg	AGG Arg	AAG Lys	GAA Glu	ATG Met	ACC Thr	AGC Ser	CTC Leu	AGG Arg	1638
		520				525					530					
GCC Ala	TTC Phe	GCT Ala	GTG Val	TAC Tyr	ACC Thr	TCC Ser	ATC Ile	TCC Ser	ATC Ile	TTC Phe	ATG Met	AAC Asn	ACA Thr	GCT Ala	ATC Ile	1686
		535			540					545					550	
CCC Pro	ATC Ile	GCT Ala	GCT Ala	GTG Val	CTC Leu	ATC Ile	ACC Thr	TTC Phe	GTG Val	GGC Gly	CAC His	GTG Val	AGC Ser	TTC Phe	TTC Phe	1734
				555					560					565		
AAA Lys	GAG Glu	TCG Ser	GAC Asp	TTC Phe	TCG Ser	CCC Pro	TCG Ser	GTG Val	GCC Ala	TTT Phe	GCC Ala	TCT Ser	CTC Leu	TCT Ser	CTC Leu	1782
			570					575					580			
TTC Phe	CAC His	ATC Ile	CTG Leu	GTG Val	ACA Thr	CCG Pro	CTG Leu	TTC Phe	CTG Leu	CTG Leu	TCT Ser	AGT Ser	GTG Val	GTT Val	CGG Arg	1830
		585					590					595				
TCC Ser	ACT Thr	GTG Val	AAG Lys	GCC Ala	CTG Leu	GTG Val	AGC Ser	GTG Val	CAA Gln	AAG Lys	CTG Leu	AGT Ser	GAG Glu	TTC Phe	CTG Leu	1878
		600				605					610					
TCC Ser	AGT Ser	GCA Ala	GAG Glu	ATC Ile	CGT Arg	GAG Glu	GAA Glu	CAG Gln	TGT Cys	GCC Ala	CCC Pro	CGA Arg	GAG Glu	CCC Pro	GCA Ala	1926
		615			620				625						630	
CCC Pro	CAA Gln	GGC Gly	CAA Gln	GCG Ala	GGC Gly	AAG Lys	TAC Tyr	CAG Gln	GCG Ala	GTG Val	CCC Pro	CTC Leu	AAG Lys	GTG Val	GTA Val	1974
				635					640					645		
AAC Asn	CGC Arg	AAG Lys	CGC Arg	CCA Pro	GCC Ala	CGA Arg	GAA Glu	GAA Glu	GTC Val	CGG Arg	GAC Asp	CTC Leu	TTG Leu	GGC Gly	CCA Pro	2022
			650				655						660			
CTG Leu	CAG Gln	AGG Arg	CTG Leu	ACT Thr	CCC Pro	AGC Ser	ACG Thr	GAT Asp	GGA Gly	GAC Asp	GCT Ala	GAC Asp	AAC Asn	TTC Phe	TGT Cys	2070
		665					670					675				
GTG Val	CAG Gln	ATC Ile	ATC Ile	GGA Gly	GGC Gly	TTC Phe	TTC Phe	ACC Thr	TGG Trp	ACC Thr	CCT Pro	GAT Asp	GGA Gly	ATC Ile	CCC Pro	2118
		680				685					690					
ACC Thr	CTG Leu	TCC Ser	AAC Asn	ATC Ile	ACC Thr	ATC Ile	CGT Arg	ATC Ile	CCC Pro	CGA Arg	GGT Gly	CAG Gln	CTG Leu	ACC Thr	ATG Met	2166
					700				705						710	
ATC Ile	GTG Val	GGG Gly	CAG Gln	GTG Val	GGC Gly	TGT Cys	GGC Gly	AAG Lys	TCC Ser	TCG Ser	CTC Leu	CTT Leu	CTG Leu	GCC Ala	ACC Thr	2214
				715					720					725		
CTG Leu	GGG Gly	GAG Glu	ATG Met	CAG Gln	AAG Lys	GTG Val	TCT Ser	GGA Gly	GCT Ala	GTG Val	TTC Phe	TGG Trp	AAC Asn	AGC Ser	CTT Leu	2262
			730					735					740			
CCA Pro	GAC Asp	AGC Ser	GAG Glu	GGG Gly	AGA Arg	AGA Arg	CCC Pro	CAG Gln	CAA Gln	CCC Pro	AGA Arg	GCG Ala	GGA Gly	GAC Asp	AGC Ser	2310
			745				750					755				

SUBSTITUTE SHEET (RULE 26)

61

GGC CGA TTC GGA TGC CAG GAG CAG AGG CCC TGT GGC TAC GCA TCT CAG Gly Arg Phe Gly Cys Gln Glu Gln Arg Pro Cys Gly Tyr Ala Ser Gln 760 765 770	2358
AAA CCA TGG CTG CTA AAT GCC ACT GTG GAG GAG AAC ATC ACC TTC GAG Lys Pro Trp Leu Leu Asn Ala Thr Val Glu Glu Asn Ile Thr Phe Glu 775 780 785 790	2406
AGT CCC TTC AAT AAG CAA CGG TAC AAG ATG GTC ATC GAA GCC TGC TCC Ser Pro Phe Asn Lys Gln Arg Tyr Lys Met Val Ile Glu Ala Cys Ser 795 800 805	2454
CTG CAG CCA GAC ATA GAC ATC CTG CCC CAT GGA GAC CAG ACT CAG ATT Leu Gln Pro Asp Ile Asp Ile Leu Pro His Gly Asp Gln Thr Gln Ile 810 815 820	2502
GGG GAA CGA GGC ATC AAC TTG AGT ACT GGT GGT CAG CGT CCA GAT CAG Gly Glu Arg Gly Ile Asn Leu Ser Thr Gly Gly Gln Arg Pro Asp Gln 825 830 835	2550
TGT AGA CCC GAG CCC TCT ACC AGC ACA CCA ATG ATT GTC TTT TTG GAT Cys Arg Pro Glu Pro Ser Thr Ser Thr Pro Met Ile Val Phe Leu Asp 840 845 850	2598
GAC CCT TTC TCG GCT CTG GAT GTC CAT CTG AGT GAC CAC CTA ATG CAG Asp Pro Phe Ser Ala Leu Asp Val His Leu Ser Asp His Leu Met Gln 855 860 865 870	2646
GCT GGC ATC CTC GAG CTG CTC CGG GAT GAC AAG AGG ACA GTG GTC TTG Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp Lys Arg Thr Val Val Leu 875 880 885	2694
GTG ACC CAC AAG CTA CAG TAC CTG CCT CAT GCT GAC TGG ATC ATT GCT Val Thr His Lys Leu Gln Tyr Leu Pro His Ala Asp Trp Ile Ile Ala 890 895 900	2742
ATG AAG GAT GGC ACC ATT CAG AGG GAG GGG ACA CTC AAG GAC TTC CAG Met Lys Asp Gly Thr Ile Gln Arg Glu Gly Thr Leu Lys Asp Phe Gln 905 910 915	2790
AGG TCT GAG TGC CAG CTC TTT GAG CAT TGG AAG ACC CTC ATG AAC CGG Arg Ser Glu Cys Gln Leu Phe Glu His Trp Lys Thr Leu Met Asn Arg 920 925 930	2838
CAG GAC CAA GAG CTG GAG AAG GAG ACA GTC ATG GAG AGA AAA GCC CCA Gln Asp Gln Glu Leu Glu Lys Glu Thr Val Met Glu Arg Lys Ala Pro 935 940 945 950	2886
GAG CCA TCT CAG GGC CTG CCC CGT GCC ATG TCC TCA AGA GAT GGC CTT Glu Pro Ser Gln Gly Leu Pro Arg Ala Met Ser Ser Arg Asp Gly Leu 955 960 965	2934
CTG CTG GAT GAG GAT GAG GAG GAA GAG GAG GCA GCC GAG AGC GAG GAA Leu Leu Asp Glu Asp Glu Glu Glu Glu Glu Ala Ala Glu Ser Glu Glu 970 975 980	2982
GAT GAC AAC TTA TCC TCT GTG CTG CAT CAG CGA GCC AAG ATC CCA TGG Asp Asp Asn Leu Ser Ser Val Leu His Gln Arg Ala Lys Ile Pro Trp 985 990 995	3030
CGA GCC TGC ACC AAG TAT TTG TCC TCT GCT GGC ATC CTG CTC CTG TCC Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala Gly Ile Leu Leu Leu Ser 1000 1005 1010	3078
CTG CTT GTC TTC TCC CAG CTG CTC AAG CAC ATG GTC TTG GTG GCC ATT Leu Leu Val Phe Ser Gln Leu Leu Lys His Met Val Leu Val Ala Ile 1015 1020 1025 1030	3126

SUBSTITUTE SHEET (RULE 26)

62

GAC TAC TGG CTG GCC AAG TGG ACG GAC AGT GCC CTG GTC CTG AGC CCC Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser Ala Leu Val Leu Ser Pro	1035	1040	1045	3174
GCC GCC AGG AAC TGC TCC CTC AGC CAG GAA TGT GCC CTG GAC CAA TCT Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu Cys Ala Leu Asp Gln Ser	1050	1055	1060	3222
GTC TAT GCC ATG GTA TTC ACC GTG CTC TGC AGC CTG GGT ATC GCG CTG Val Tyr Ala Met Val Phe Thr Val Leu Cys Ser Leu Gly Ile Ala Leu	1065	1070	1075	3270
TGC CTT GTC ACC TCT GTC ACT GTG GAG TGG ACG GGA CTG AAG GTG GCC Cys Leu Val Thr Ser Val Thr Val Glu Trp Thr Gly Leu Lys Val Ala	1080	1085	1090	3318
AAG AGG CTG CAT CGC AGC CTG CTC AAC CGT ATC ATC CTG GCT CCC ATG Lys Arg Leu His Arg Ser Leu Leu Asn Arg Ile Ile Leu Ala Pro Met	1095	1100	1105	3366
AGG TTC TTT GAG ACC ACG CCC CTG GGG AGT ATC CTG AAC AGA TTT TCA Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser Ile Leu Asn Arg Phe Ser	1115	1120	1125	3414
TCT GAC TGT AAC ACC ATT GAC CAG CAT ATC CCG TCC ACG CTG GAG TGC Ser Asp Cys Asn Thr Ile Asp Gln His Ile Pro Ser Thr Leu Glu Cys	1130	1135	1140	3462
CTG AGC AGA TCC ACC TTA CTC TGT GTC TCC GCC CTG GCT GTC ATC TCC Leu Ser Arg Ser Thr Leu Leu Cys Val Ser Ala Leu Ala Val Ile Ser	1145	1150	1155	3510
TAC GTC ACG CCT GTG TTC CTA GTG GCC CTC TTA CCC CTC GCC GTC GTG Tyr Val Thr Pro Val Phe Leu Val Ala Leu Leu Pro Leu Ala Val Val	1160	1165	1170	3558
TGC TAC TTC ATC CAG AAG TAC TTC CGA GTG GCG TCC AGG GAC CTG CAG Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val Ala Ser Arg Asp Leu Gln	1175	1180	1185	3606
CAG CTG GAC GAC ACA ACA CAG CTC CCT CTG CTC TCA CAC TTT GCT GAA Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu Leu Ser His Phe Ala Glu	1195	1200	1205	3654
ACT GTG GAA GGA CTC ACC ACC ATC CGT GCC TTC AGG TAC GAG GCC CGG Thr Val Glu Gly Leu Thr Thr Ile Arg Ala Phe Arg Tyr Glu Ala Arg	1210	1215	1220	3702
TTC CAG CAG AAG CTC CTA GAG TAC ACC GAC TCC AAC AAC ATT GCC TCT Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp Ser Asn Asn Ile Ala Ser	1225	1230	1235	3750
CTC TTC CTC ACA GCA GCC AAC AGG TGG CTG GAA GTC CGC ATG GAG TAC Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu Glu Val Arg Met Glu Tyr	1240	1245	1250	3798
ATC GGA GCA TGC GTG GTA CTC ATC GCC GCT GCC ACC TCC ATC TCC AAC Ile Gly Ala Cys Val Leu Leu Ile Ala Ala Thr Ser Ile Ser Asn	1255	1260	1265	3846
TCC CTA CAC AGG GAG CTC TCA GCC GGC CTA GTA GGC CTG GGC CTC ACC Ser Leu His Arg Glu Leu Ser Ala Gly Leu Val Gly Leu Gly Leu Thr	1275	1280	1285	3894
TAT GCC TTG ATG ATT GGG ATC TGC GGC CGC ACA GGC AGT GGA AAA TCC Tyr Ala Leu Met Ile Gly Ile Cys Gly Arg Thr Gly Ser Gly Lys Ser	1290	1295	1300	3942

SUBSTITUTE SHEET (RULE 26)

63

TCC TTC TCT CTC GCC TTT TTC CGA ATG GTG GAT ATG TTT GAA GGG CGT Ser Phe Ser Leu Ala Phe Phe Arg Met Val Asp Met Phe Glu Gly Arg 1305 1310 1315	3990
ATC ATC ATC GAT GGC ATT GAC ATC GCC AAG CTG CCG CTG CAC ACG CTC Ile Ile Ile Asp Gly Ile Asp Ile Ala Lys Leu Pro Leu His Thr Leu 1320 1325 1330	4038
CGC TCA CGC CTG TCT ATC ATC CTA CAG GAC CCT GTT CTC TTC AGT GGT Arg Ser Arg Leu Ser Ile Ile Leu Gln Asp Pro Val Leu Phe Ser Gly 1335 1340 1345 1350	4086
ACC ATC AGA TTC AAC CTG GAC CCA GAG AAG AAA TGC TCA GAC AGC ACC Thr Ile Arg Phe Asn Leu Asp Pro Glu Lys Lys Cys Ser Asp Ser Thr 1355 1360 1365	4134
CTG TGG GAG GCT CTG GAG ATC GCT CAG CTG AAG CTG GTG GTG AAG GCC Leu Trp Glu Ala Leu Glu Ile Ala Gln Leu Lys Leu Val Val Lys Ala 1370 1375 1380	4182
CTG CCA GGA GGC CTG GAT GCC ATC ATC ACG GAA GGA GGG GAG AAT TTT Leu Pro Gly Gly Leu Asp Ala Ile Ile Thr Glu Gly Gly Glu Asn Phe 1385 1390 1395	4230
AGC CAG GGC CAG AGG CAG CTG TTC TGC CTG GCC CGG GCC TTT GTG AGG Ser Gln Gly Gln Arg Gln Leu Phe Cys Leu Ala Arg Ala Phe Val Arg 1400 1405 1410	4278
AAG ACC AGC ATC TTC ATC ATG GAT GAA GCA ACT GCC TCC ATC GAC ATG Lys Thr Ser Ile Phe Ile Met Asp Glu Ala Thr Ala Ser Ile Asp Met 1415 1420 1425 1430	4326
GCT ACG GAA AAT ATC CTC CAG AAG GTG GTG ATG ACA GCC TTC GCA GAC Ala Thr Glu Asn Ile Leu Gln Lys Val Val Met Thr Ala Phe Ala Asp 1435 1440 1445	4374
CGC ACC GTG GTC ACC ATC GCG CAC CGC GTG CAC ACC ATC CTG AGT GCA Arg Thr Val Val Thr Ile Ala His Arg Val His Thr Ile Leu Ser Ala 1450 1455 1460	4422
GAC CTA GTG ATG GTC CTG AAG AGG GGC GCG ATC CTG GAG TTC GAC AAG Asp Leu Val Met Val Leu Lys Arg Gly Ala Ile Leu Glu Phe Asp Lys 1465 1470 1475	4470
CCG GAA AAG CTT CTC AGC CAG AAG GAC AGC GTC TTT GCC TCC TTT GTC Pro Glu Lys Leu Leu Ser Gln Lys Asp Ser Val Phe Ala Ser Phe Val 1480 1485 1490	4518
CGC GCG GAC AAA TGACCAGCCA GCGCCAAAGT GCCACCCAC ACCTCACCTG Arg Ala Asp Lys	4570
CTTGCCATGG ATTTCTTACT GTAAATCACT TGTAATAAAA GAAACTAATT CTTTGCTAAA	4630
AAAAA	4635

## (6) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1498 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

SUBSTITUTE SHEET (RULE 26)

64

Met Pro Leu Ala Phe Cys Gly Thr Glu Asn His Ser Ala Ala Tyr Arg  
 1 5 10 15  
 Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn  
 20 25 30  
 Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe  
 35 40 45  
 Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr  
 50 55 60  
 Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Ile  
 65 70 75 80  
 Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp  
 85 90 95  
 Gly Val Thr Glu Ser Arg His Leu His Leu Tyr Met Pro Ala Gly Met  
 100 105 110  
 Ala Phe Met Ala Ala Ile Thr Ser Val Val Tyr Tyr His Asn Ile Glu  
 115 120 125  
 Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr  
 130 135 140  
 Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Tyr Asp His  
 145 150 155 160  
 Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val  
 165 170 175  
 Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val  
 180 185 190  
 Arg Arg Tyr Val Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu  
 195 200 205  
 Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu  
 210 215 220  
 Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala  
 225 230 235 240  
 His Lys Lys Pro Ile Asp Leu Arg Ala Ile Gly Lys Leu Pro Ile Ala  
 245 250 255  
 Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Leu Ala Phe Asp Ala  
 260 265 270  
 Gln Ala Arg Lys Asp Thr Gln Ser Gln Gln Gly Ala Arg Ala Ile Trp  
 275 280 285  
 Arg Ala Leu Cys His Ala Phe Gly Arg Arg Leu Val Leu Ser Ser Thr  
 290 295 300  
 Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile  
 305 310 315 320  
 Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn His Val Phe Gln Pro  
 325 330 335  
 Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu  
 340 345 350  
 Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu

SUBSTITUTE SHEET (RULE 26)

65

355	360	365
Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly 370 375 380		
Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met 385 390 395 400		
His Leu Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln 405 410 415		
Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe 420 425 430		
Phe Leu Cys Pro Asn Leu Trp Ala Met Pro Val Gln Ile Ile Val Gly 435 440 445		
Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala 450 455 460		
Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys 465 470 475 480		
Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu Tyr Ser Asn Glu Arg Leu 485 490 495		
Lys Gln Thr Asn Glu Met Leu Arg Gly Ile Lys Leu Leu Lys Leu Tyr 500 505 510		
Ala Trp Glu Asn Ile Phe Cys Ser Arg Val Glu Lys Thr Arg Arg Lys 515 520 525		
Glu Met Thr Ser Leu Arg Ala Phe Ala Val Tyr Thr Ser Ile Ser Ile 530 535 540		
Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val 545 550 555 560		
Gly His Val Ser Phe Phe Lys Glu Ser Asp Phe Ser Pro Ser Val Ala 565 570 575		
Phe Ala Ser Leu Ser Leu Phe His Ile Leu Val Thr Pro Leu Phe Leu 580 585 590		
Leu Ser Ser Val Val Arg Ser Thr Val Lys Ala Leu Val Ser Val Gln 595 600 605		
Lys Leu Ser Glu Phe Leu Ser Ser Ala Glu Ile Arg Glu Glu Gln Cys 610 615 620		
Ala Pro Arg Glu Pro Ala Pro Gln Gly Gln Ala Gly Lys Tyr Gln Ala 625 630 635 640		
Val Pro Leu Lys Val Val Asn Arg Lys Arg Pro Ala Arg Glu Glu Val 645 650 655		
Arg Asp Leu Leu Gly Pro Leu Gln Arg Leu Thr Pro Ser Thr Asp Gly 660 665 670		
Asp Ala Asp Asn Phe Cys Val Gln Ile Ile Gly Gly Phe Phe Thr Trp 675 680 685		
Thr Pro Asp Gly Ile Pro Thr Leu Ser Asn Ile Thr Ile Arg Ile Pro 690 695 700		
Arg Gly Gln Leu Thr Met Ile Val Gly Gln Val Gly Cys Gly Lys Ser 705 710 715 720		

SUBSTITUTE SHEET (RULE 26)

66

Ser Leu Leu Leu Ala Thr Leu Gly Glu Met Gln Lys Val Ser Gly Ala  
 725 730 735  
 Val Phe Trp Asn Ser Leu Pro Asp Ser Glu Gly Arg Arg Pro Gln Gln  
 740 745 750  
 Pro Arg Ala Gly Asp Ser Gly Arg Phe Gly Cys Gln Glu Gln Arg Pro  
 755 760 765  
 Cys Gly Tyr Ala Ser Gln Lys Pro Trp Leu Leu Asn Ala Thr Val Glu  
 770 775 780  
 Glu Asn Ile Thr Phe Glu Ser Pro Phe Asn Lys Gln Arg Tyr Lys Met  
 785 790 795 800  
 Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile Asp Ile Leu Pro His  
 805 810 815  
 Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile Asn Leu Ser Thr Gly  
 820 825 830  
 Gly Gln Arg Pro Asp Gln Cys Arg Pro Glu Pro Ser Thr Ser Thr Pro  
 835 840 845  
 Met Ile Val Phe Leu Asp Asp Pro Phe Ser Ala Leu Asp Val His Leu  
 850 855 860  
 Ser Asp His Leu Met Gln Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp  
 865 870 875 880  
 Lys Arg Thr Val Val Leu Val Thr His Lys Leu Gln Tyr Leu Pro His  
 885 890 895  
 Ala Asp Trp Ile Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly  
 900 905 910  
 Thr Leu Lys Asp Phe Gln Arg Ser Glu Cys Gln Leu Phe Glu His Trp  
 915 920 925  
 Lys Thr Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val  
 930 935 940  
 Met Glu Arg Lys Ala Pro Glu Pro Ser Gln Gly Leu Pro Arg Ala Met  
 945 950 955 960  
 Ser Ser Arg Asp Gly Leu Leu Leu Asp Glu Asp Glu Glu Glu Glu  
 965 970 975  
 Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Val Leu His Gln  
 980 985 990  
 Arg Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala  
 995 1000 1005  
 Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His  
 1010 1015 1020  
 Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser  
 1025 1030 1035 1040  
 Ala Leu Val Leu Ser Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu  
 1045 1050 1055  
 Cys Ala Leu Asp Gln Ser Val Tyr Ala Met Val Phe Thr Val Leu Cys  
 1060 1065 1070  
 Ser Leu Gly Ile Ala Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp

SUBSTITUTE SHEET (RULE 26)



1075	1080	1085
Thr Gly Leu Lys Val	Ala Lys Arg Leu His Arg	Ser Leu Leu Asn Arg
1090	1095	1100
Ile Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser		
1105	1110	1115 1120
Ile Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile		
	1125 1130	1135
Pro Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser		
	1140 1145	1150
Ala Leu Ala Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu		
	1155 1160	1165
Leu Pro Leu Ala Val Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val		
	1170 1175	1180
Ala Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu		
1185	1190	1195 1200
Leu Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala		
	1205 1210	1215
Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp		
	1220 1225	1230
Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu		
	1235 1240	1245
Glu Val Arg Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala		
	1250 1255	1260
Ala Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu		
1265	1270	1275 1280
Val Gly Leu Gly Leu Thr Tyr Ala Leu Met Ile Gly Ile Cys Gly Arg		
	1285 1290	1295
Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe Phe Arg Met Val		
	1300 1305	1310
Asp Met Phe Glu Gly Arg Ile Ile Ile Asp Gly Ile Asp Ile Ala Lys		
	1315 1320	1325
Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile Ile Leu Gln Asp		
	1330 1335	1340
Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu Asp Pro Glu Lys		
1345	1350	1355 1360
Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu Ile Ala Gln Leu		
	1365 1370	1375
Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp Ala Ile Ile Thr		
	1380 1385	1390
Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln Leu Phe Cys Leu		
	1395 1400	1405
Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile Met Asp Glu Ala		
	1410 1415	1420
Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu Gln Lys Val Val		
1425	1430	1435 1440

SUBSTITUTE SHEET (RULE 26)

68

Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile Ala His Arg Val  
 1445 1450 1455

His Thr Ile Leu Ser Ala Asp Leu Val Met Val Leu Lys Arg Gly Ala  
 1460 1465 1470

Ile Leu Glu Phe Asp Lys Pro Glu Lys Leu Leu Ser Gln Lys Asp Ser  
 1475 1480 1485

Val Phe Ala Ser Phe Val Arg Ala Asp Lys  
 1490 1495

## (7) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 4625 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS  
 (B) LOCATION: 25..4521

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

AGCCGAGCCC GTGCGCGCGC CGCC ATG CCC TTG GCC TTC TGC GGT ACC GAG	51
Met Pro Leu Ala Phe Cys Gly Thr Glu	
1 5	
AAC CAC TCG GCC GCC TAC CGG GTG GAC CAG GGC GTC CTC AAC AAC GGC	99
Asn His Ser Ala Ala Tyr Arg Val Asp Gln Gly Val Leu Asn Asn Gly	
10 15 20 25	
TGC TTC GTG GAC GCG CTC AAC GTG GTG CCG CAC GTT TTC CTG CTC TTC	147
Cys Phe Val Asp Ala Leu Asn Val Val Pro His Val Phe Leu Leu Phe	
30 35 40	
ATC ACC TTC CCC ATC CTC TTC ATC GGA TGG GGC AGC CAG AGC TCC AAG	195
Ile Thr Phe Pro Ile Leu Phe Ile Gly Trp Gly Ser Gln Ser Ser Lys	
45 50 55	
GTG CAC ATC CAC CAC AGC ACC TGG CTG CAC TTT CCA GGG CAC AAC CTG	243
Val His Ile His His Ser Thr Trp Leu His Phe Pro Gly His Asn Leu	
60 65 70	
CGC TGG ATC CTT ACC TTC ATT TTG CTC TTC GTC CTT GTG TGT GAG ATC	291
Arg Trp Ile Leu Thr Phe Ile Leu Leu Phe Val Leu Val Cys Glu Ile	
75 80 85	
GCT GAG GGC ATC CTG TCT GAT GGG GTG ACA GAA TCC CGC CAC CTC CAC	339
Ala Glu Gly Ile Leu Ser Asp Gly Val Thr Glu Ser Arg His Leu His	
90 95 100 105	
CTG TAC ATG CCA GCC GGG ATG GCG TTC ATG GCT GCC ATC ACC TCT GTA	387
Leu Tyr Met Pro Ala Gly Met Ala Phe Met Ala Ala Ile Thr Ser Val	
110 115 120	

SUBSTITUTE SHEET (RULE 26)

GTC TAC TAT CAT AAC ATC GAG ACC TCC AAC TTC CCC AAG CTT TTG ATC Val Tyr Tyr His Asn Ile Glu Thr Ser Asn Phe Pro Lys Leu Leu Ile 125 130 135	435
GCT CTG CTC ATC TAT TGG ACC CTG GCC TTC ATC ACG AAG ACC ATC AAG Ala Leu Leu Ile Tyr Trp Thr Leu Ala Phe Ile Thr Lys Thr Ile Lys 140 145 150	483
TTT GTC AAG TTC TAT GAC CAC GCC ATC GGC TTC TCC CAG CTG CGC TTC Phe Val Lys Phe Tyr Asp His Ala Ile Gly Phe Ser Gln Leu Arg Phe 155 160 165	531
TGC CTC ACG GGG CTT CTG GTG ATC CTG TAT GGG ATG TTG CTG CTT GTG Cys Leu Thr Gly Leu Leu Val Ile Leu Tyr Gly Met Leu Leu Leu Val 170 175 180 185	579
GAG GTC AAC GTC ATC AGA GTG AGG AGG TAC ATC TTC TTC AAG ACG CCA Glu Val Asn Val Ile Arg Val Arg Arg Tyr Ile Phe Phe Lys Thr Pro 190 195 200	627
CGG GAG GTG AAG CCC CCT GAG GAC CTG CAG GAC CTG GGT GTG CGC TTT Arg Glu Val Lys Pro Pro Glu Asp Leu Gln Asp Leu Gly Val Arg Phe 205 210 215	675
CTG CAG CCC TTC GTT AAC CTG CTG TCA AAG GGG ACC TAT TGG TGG ATG Leu Gln Pro Phe Val Asn Leu Leu Ser Lys Gly Thr Tyr Trp Trp Met 220 225 230	723
AAT GCC TTC ATC AAG ACG GCC CAC AAG AAG CCC ATC GAC CTG CGG GCC Asn Ala Phe Ile Lys Thr Ala His Lys Lys Pro Ile Asp Leu Arg Ala 235 240 245	771
ATC GCG AAG CTG CCC ATC GCC ATG AGA GCC CTC ACC AAC TAT CAG CGC Ile Ala Lys Leu Pro Ile Ala Met Arg Ala Leu Thr Asn Tyr Gln Arg 250 255 260 265	819
CTC TGC GTG GCC TTC GAT GCT CAG GCG CGG AAG GAC ACA CAG AGC CCA Leu Cys Val Ala Phe Asp Ala Gln Ala Arg Lys Asp Thr Gln Ser Pro 270 275 280	867
CAG GGT GCC CGG GCC ATC TGG AGG GCT CTA TGC CAT GCC TTT GGG AGA Gln Gly Ala Arg Ala Ile Trp Arg Ala Leu Cys His Ala Phe Gly Arg 285 290 295	915
CGC CTG ATC CTC AGC AGC ACA TTC CGC ATC CTG GCT GAC CTG TTG GGC Arg Leu Ile Leu Ser Ser Thr Phe Arg Ile Leu Ala Asp Leu Leu Gly 300 305 310	963
TTC GCT GGA CCA CTC TGC ATC TTT GGG ATC GTG GAC CAC CTG GGG AAG Phe Ala Gly Pro Leu Cys Ile Phe Gly Ile Val Asp His Leu Gly Lys 315 320 325	1011
GAG AAC CAC GTC TTC CAG CCC AAG ACA CAG TTT CTC GGG GTT TAC TTC Glu Asn His Val Phe Gln Pro Lys Thr Gln Phe Leu Gly Val Tyr Phe 330 335 340 345	1059
GTC TCT TCT CAA GAG TTC CTT GGC AAT GCC TAC GTC TTG GCC GTG CTT Val Ser Ser Gln Glu Phe Leu Gly Asn Ala Tyr Val Leu Ala Val Leu 350 355 360	1107
CTG TTC CTT GCC CTG CTA CTG CAA AGG ACA TTC CTG CAA GCC TCA TAC Leu Phe Leu Ala Leu Leu Leu Gln Arg Thr Phe Leu Gln Ala Ser Tyr 365 370 375	1155
TAC GTC GCC ATT GAA ACT GGA ATT AAC CTG AGA GGA GCA ATC CAG ACC Tyr Val Ala Ile Glu Thr Gly Ile Asn Leu Arg Gly Ala Ile Gln Thr 380 385 390	1203

SUBSTITUTE SHEET (RULE 26)

70

AAG Lys	ATT Ile	TAC Tyr	AAT Asn	AAA Lys	ATC Ile	ATG Met	CAC His	ATG Met	TCC Ser	ACC Thr	TCC Ser	AAC Asn	CTG Leu	TCA Ser	ATG Met	1251
395						400				405						
GGG Gly	GAA Glu	ATG Met	ACT Thr	GCT Ala	GGG Gly	CAG Gln	ATC Ile	TGC Cys	AAC Asn	CTG Leu	GTG Val	GCC Ala	ATC Ile	GAC Asp	ACA Thr	1299
410					415					420					425	
AAC Asn	CAG Gln	CTC Leu	ATG Met	TGG Trp	TTC Phe	TTC Phe	TTT Phe	CTG Leu	TGC Cys	CCA Pro	AAC Asn	CTC Leu	TGG Trp	ACG Thr	ATG Met	1347
				430					435					440		
CCA Pro	GTA Val	CAG Gln	ATC Ile	ATT Ile	GTG Val	GGC Gly	GTG Val	ATC Ile	CTT Leu	CTC Leu	TAC Tyr	TAC Tyr	ATC Ile	CTT Leu	GGG Gly	1395
				445				450					455			
GTC Val	AGT Ser	GCC Ala	TTG Leu	ATT Ile	GGA Gly	GCA Ala	GCT Ala	ATC Val	ATT Ile	CTG Leu	CTG Leu	GCT Ala	CCT Pro	GTA Val		1443
		460					465					470				
CAG Gln	TAC Tyr	TTT Phe	GTG Val	GCC Ala	ACC Thr	AAG Lys	CTC Leu	TCC Ser	CAG Gln	GCA Ala	CAG Gln	CGG Arg	ACG Thr	ACC Thr	TTG Leu	1491
	475					480					485					
GAG Glu	CAC His	TCC Ser	AAC Asn	GAG Glu	AGG Arg	CTG Leu	AAG Lys	CAG Gln	ACC Thr	AAC Asn	GAG Glu	ATG Met	CTC Leu	CGG Arg	GGC Gly	1539
490					495					500					505	
ATG Met	AAG Lys	CTG Leu	CTC Leu	AAA Lys	CTG Leu	TAT Tyr	GCG Ala	TGG Trp	GAG Glu	AGC Ser	ATC Ile	TTC Phe	TGC Cys	TCC Ser	AGG Arg	1587
				510					515					520		
GTG Val	GAG Glu	GTG Val	ACT Thr	CGC Arg	AGG Arg	AAG Lys	GAG Glu	ATG Met	ACC Thr	AGC Ser	CTG Leu	AGG Arg	GCG Ala	TTT Phe	GCT Ala	1635
			525					530					535			
GTC Val	TAC Tyr	ACT Thr	TCC Ser	ATC Ile	TCC Ser	ATC Ile	TTC Phe	ATG Met	AAC Asn	ACA Thr	GCC Ala	ATC Ile	CCC Pro	ATT Ile	GCT Ala	1683
			540				545					550				
GCC Ala	GTG Val	CTC Leu	ATC Ile	ACC Thr	TTC Phe	GTG Val	GGC Gly	CAC His	GTG Val	AGC Ser	TTC Phe	TTC Phe	AAA Lys	GAG Glu	TCG Ser	1731
		555				560					565					
GAC Asp	TTG Leu	TCA Ser	CCC Pro	TCG Ser	GTG Val	GCC Ala	TTT Phe	GCC Ala	TCC Ser	CTC Leu	TCT Ser	CTC Leu	TTC Phe	CAC His	ATC Ile	1779
570					575					580				585		
CTG Leu	GTG Val	ACT Thr	CCA Pro	CTG Leu	TTC Phe	CTG Leu	TCT Leu	AGC Ser	GTG Val	GTT Val	CGG Arg	TCC Ser	ACT Thr	GTC Val		1827
			590					595					600			
AAA Lys	GCC Ala	CTG Leu	GTG Val	AGC Ser	GTG Val	CAA Gln	AAA Lys	CTG Leu	AGC Ser	GAG Glu	TTC Phe	CTG Leu	TCT Ser	AGT Ser	GCA Ala	1875
		605					610						615			
GAG Glu	ATC Ile	CGT Arg	GAG Glu	GAG Glu	CAG Gln	TGT Cys	GCC Ala	CCC Pro	CGA Arg	GAG Glu	CCT Pro	GCA Ala	CCC Pro	CAA Gln	GGC Gly	1923
		620					625					630				
CAA Gln	GCC Ala	GGC Gly	AAG Lys	TAC Tyr	CAG Gln	GCA Ala	GTG Val	CCC Pro	CTC Leu	AAG Lys	GTT Val	GTG Val	AAC Asn	CGC Arg	AAA Lys	1971
		635				640					645					
CGC Arg	CCA Pro	GCC Ala	CGG Arg	GAA Glu	GAG Glu	GTG Val	CGG Arg	GAC Asp	CTC Leu	CTG Leu	GGC Gly	CCA Pro	CTG Leu	CAG Gln	AGG Arg	2019
					655					660					665	

SUBSTITUTE SHEET (RULE 26)

CTG GCC CCT AGC ATG GAC GGG GAT GCT GAC AAC TTC TGT GTC CAG ATC Leu Ala Pro Ser Met Asp Gly Asp Ala Asp Asn Phe Cys Val Gln Ile 670 675 680	2067
ATC GGA GGC TTC TTC ACC TGG ACC CCT GAT GGA ATC CCC ACT CTG TCC Ile Gly Gly Phe Phe Thr Trp Thr Pro Asp Gly Ile Pro Thr Leu Ser 685 690 695	2115
AAC ATC ACC ATC CGT ATT CCC CGA GGT CAG CTA ACC ATG ATT GTG GGG Asn Ile Thr Ile Arg Ile Pro Arg Gly Gln Leu Thr Met Ile Val Gly 700 705 710	2163
CAG GTG GGC TGC GGC AAG TCC TCG CTC CTC CTC GCC ACC CTG GGG GAG Gln Val Gly Cys Gly Lys Ser Ser Leu Leu Leu Ala Thr Leu Gly Glu 715 720 725	2211
ATG CAG AAG GTG TCG GGG GCC GTC TTC TGG AAC AGC AAC CTT CCG GAC Met Gln Lys Val Ser Gly Ala Val Phe Trp Asn Ser Asn Leu Pro Asp 730 735 740 745	2259
AGC GAG GGG AGA GGA CCC CAG CAG CCC AGA GCG GGA GAC AGC AGC TGG Ser Glu Gly Arg Gly Pro Gln Gln Pro Arg Ala Gly Asp Ser Ser Trp 750 755 760	2307
CTC GGA TAT CAG GAG CAG AGG CCC CGT GGC TAC GCA TCT CAG AAA CCA Leu Gly Tyr Gln Glu Gln Arg Pro Arg Gly Tyr Ala Ser Gln Lys Pro 765 770 775	2355
TGG CTG CTA AAC GCC ACC GTG GAA GAG AAC ATC ACC TTC GAG AGT CCC Trp Leu Leu Asn Ala Thr Val Glu Glu Asn Ile Thr Phe Glu Ser Pro 780 785 790	2403
TTC AAT CCG CAG CGG TAC AAG ATG GTC ATC GAA GCC TGC TCC CTG CAG Phe Asn Pro Gln Arg Tyr Lys Met Val Ile Glu Ala Cys Ser Leu Gln 795 800 805	2451
CCG GAC ATA GAC ATC CTG CCC CAC GGA GAC CAG ACT CAG ATT GGG GAA Pro Asp Ile Asp Ile Leu Pro His Gly Asp Gln Thr Gln Ile Gly Glu 810 815 820 825	2499
CGG GGC ATC AAC CTG TCT GGT GGT CAG CGT CCA GAT CAG TGT GGT CCA Arg Gly Ile Asn Leu Ser Gly Gly Gln Arg Pro Asp Gln Cys Gly Pro 830 835 840	2547
GAG CCC TCT ACC AGC AGA CCA ATG TTC GTC TTC TTG GAT GAC CCC TTC Glu Pro Ser Thr Ser Arg Pro Met Phe Val Phe Leu Asp Asp Pro Phe 845 850 855	2595
TCA GCT TTG GAT GTC CAT CTG AGT GAC CAC CTG ATG CAG GCC GGC ATC Ser Ala Leu Asp Val His Leu Ser Asp His Leu Met Gln Ala Gly Ile 860 865 870	2643
CTT GAG CTG CTC CGG GAT GAC AAG AGG ACA GTG GTC TTG GTG ACC CAC Leu Glu Leu Leu Arg Asp Asp Lys Arg Thr Val Val Leu Val Thr His 875 880 885	2691
AAG CTA CAG TAT CTG CCT CAT GCA GAC TGG ATC ATT GCC ATG AAG GAT Lys Leu Gln Tyr Leu Pro His Ala Asp Trp Ile Ile Ala Met Lys Asp 890 895 900 905	2739
GGG ACC ATT CAG AGG GAA GGG ACG CTC AAG GAC TTC CAG AGG TCC GAG Gly Thr Ile Gln Arg Glu Gly Thr Leu Lys Asp Phe Gln Arg Ser Glu 910 915 920	2787
TGC CAG CTC TTT GAG CAC TGG AAG ACC CTC ATG AAC CGG CAG GAC CAA Cys Gln Leu Phe Glu His Trp Lys Thr Leu Met Asn Arg Gln Asp Gln 925 930 935	2835

SUBSTITUTE SHEET (RULE 26)

GAG CTG GAG AAG GAG ACA GTC ATG GAG AGG AAA GCC TCA GAG CCA TCT Glu Leu Glu Lys Glu Thr Val Met Glu Arg Lys Ala Ser Glu Pro Ser	2883
940 945 950	
CAG GGC CTG CCC CGT GCC ATG TCC TCC AGA GAC GGC CTT CTG CTG GAT Gln Gly Leu Pro Arg Ala Met Ser Ser Arg Asp Gly Leu Leu Leu Asp	2931
955 960 965	
GAG GAA GAG GAG GAA GAG GAG GCA GCC GAA AGC GAG GAA GAT GAC AAC Glu Glu Glu Glu Glu Glu Glu Ala Ala Glu Ser Glu Glu Asp Asp Asn	2979
970 975 980 985	
TTA TCT TCA GTG CTG CAT CAG CGA GCT AAG ATC CCC TGG CGA GCC TGC Leu Ser Ser Val Leu His Gln Arg Ala Lys Ile Pro Trp Arg Ala Cys	3027
990 995 1000	
ACT AAG TAT CTG TCC TCT GCT GGC ATT CTG CTC CTG TCC CTG CTT GTC Thr Lys Tyr Leu Ser Ser Ala Gly Ile Leu Leu Leu Ser Leu Leu Val	3075
1005 1010 1015	
TTC TCC CAG CTG CTC AAG CAC ATG GTC TTG GTG GCC ATT GAT TAT TGG Phe Ser Gln Leu Leu Lys His Met Val Leu Val Ala Ile Asp Tyr Trp	3123
1020 1025 1030	
CTG GCC AAG TGG ACG GAC AGT GCC CTG GTC CTG AGC CCC GCT GCC AGG Leu Ala Lys Trp Thr Asp Ser Ala Leu Val Leu Ser Pro Ala Ala Arg	3171
1035 1040 1045	
AAC TGT TCG CTC AGC CAG GAA TGT GAC CTG GAC CAG TCT GTC TAT GCC Asn Cys Ser Leu Ser Gln Glu Cys Asp Leu Asp Gln Ser Val Tyr Ala	3219
1050 1055 1060 1065	
ATG GTA TTC ACC TTG CTC TGC AGC CTG GGT ATC GTG CTG TGC CTG GTC Met Val Phe Thr Leu Leu Cys Ser Leu Gly Ile Val Leu Cys Leu Val	3267
1070 1075 1080	
ACC TCT GTC ACT GTG GAG TGG ACG GGA CTG AAG GTG GCC AAG AGG CTA Thr Ser Val Thr Val Glu Trp Thr Gly Leu Lys Val Ala Lys Arg Leu	3315
1085 1090 1095	
CAC CGC AGC CTG CTC AAC CGC ATC ATC CTG GCC CCC ATG AGG TTC TTT His Arg Ser Leu Leu Asn Arg Ile Ile Leu Ala Pro Met Arg Phe Phe	3363
1100 1105 1110	
GAG ACC ACA CCC CTC GGG AGT ATC CTG AAC AGA TTT TCA TCC GAC TGT Glu Thr Thr Pro Leu Gly Ser Ile Leu Asn Arg Phe Ser Ser Asp Cys	3411
1115 1120 1125	
AAC ACC ATT GAC CAG CAC ATC CCA TCC ACG CTG GAG TGT CTG AGC CGG Asn Thr Ile Asp Gln His Ile Pro Ser Thr Leu Glu Cys Leu Ser Arg	3459
1130 1135 1140 1145	
TCC ACC CTG CTG TGT GTC TCC GCC CTG ACT GTC ATC TCC TAT GTC ACA Ser Thr Leu Leu Cys Val Ser Ala Leu Thr Val Ile Ser Tyr Val Thr	3507
1150 1155 1160	
CCC GTG TTC CTC GTG GCC CTC TTA CCC CTA GCT GTT GTG TGC TAC TTC Pro Val Phe Leu Val Ala Leu Leu Pro Leu Ala Val Val Cys Tyr Phe	3555
1165 1170 1175	
ATT CAG AAG TAC TTC CGA GTG GCA TCC AGG GAC CTG CAG CAG CTG GAC Ile Gln Lys Tyr Phe Arg Val Ala Ser Arg Asp Leu Gln Gln Leu Asp	3603
1180 1185 1190	
GAC ACG ACG CAG CTC CCG CTC GTC TCA CAC TTT GCT GAA ACT GTG GAG Asp Thr Thr Gln Leu Pro Leu Val Ser His Phe Ala Glu Thr Val Glu	3651
1195 1200 1205	

SUBSTITUTE SHEET (RULE 26)

GGA CTC ACC ACC ATC CGT GCC TTC AGG TAC GAG GCC CGG TTC CAG CAG Gly Leu Thr Thr Ile Arg Ala Phe Arg Tyr Glu Ala Arg Phe Gln Gln 1210 1215 1220 1225	3699
AAG CTT CTA GAA TAT ACC GAC TCC AAC AAC ATC GCC TCC CTC TTC CTC Lys Leu Leu Glu Tyr Thr Asp Ser Asn Asn Ile Ala Ser Leu Phe Leu 1230 1235 1240	3747
ACG GCA GCC AAC AGA TGG CTG GAA GTC TGC ATG GAG TAC ATC GGA GCG Thr Ala Ala Asn Arg Trp Leu Glu Val Cys Met Glu Tyr Ile Gly Ala 1245 1250 1255	3795
TGC GTG GTA CTC ATT GCG GCT GCC ACC TCC ATC TCC AAC TCC CTG CAC Cys Val Val Leu Ile Ala Ala Ala Thr Ser Ile Ser Asn Ser Leu His 1260 1265 1270	3843
AGG GAA CTT TCT GCT GGC CTG GTG GGC CTG GGC CTC ACC TAT GCC TTG Arg Glu Leu Ser Ala Gly Leu Val Gly Leu Gly Leu Thr Tyr Ala Leu 1275 1280 1285	3891
ATG ATC GGG ATC TGC GGC CGC ACA GCG AGC GGG AAG TCC TCC TTC TCC Met Ile Gly Ile Cys Gly Arg Thr Ala Ser Gly Lys Ser Ser Phe Ser 1290 1295 1300 1305	3939
CTG GCC TTT TTC CGA ATG GTG GAC ATG TTT GAA GGA CGC ATC ATC ATT Leu Ala Phe Phe Arg Met Val Asp Met Phe Glu Gly Arg Ile Ile Ile 1310 1315 1320	3987
GAT GGC ATC GAC ATC GCC AAG CTG CCA CTT CAC ACG CTG CGC TCA CGC Asp Gly Ile Asp Ile Ala Lys Leu Pro Leu His Thr Leu Arg Ser Arg 1325 1330 1335	4035
CTG TCC ATC ATC CTA CAG GAC CCC GTC CTC TTC AGC GGC ACG ATC AGA Leu Ser Ile Ile Leu Gln Asp Pro Val Leu Phe Ser Gly Thr Ile Arg 1340 1345 1350	4083
TTC AAC CTG GAC CCC GAG AAG AAA TGC TCA GAC AGC ACA CTG TGG GAG Phe Asn Leu Asp Pro Glu Lys Lys Cys Ser Asp Ser Thr Leu Trp Glu 1355 1360 1365	4131
GCC CTG GAG ATC GCC CAG CTG AAG CTG GTA GTG AAG GCA CTG CCA GGA Ala Leu Glu Ile Ala Gln Leu Lys Leu Val Val Lys Ala Leu Pro Gly 1370 1375 1380 1385	4179
GGC CTA GAT GCC ATC ATC ACA GAA GGA GGG GAG AAT TTT AGC CAG GGC Gly Leu Asp Ala Ile Ile Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly 1390 1395 1400	4227
CAG AGG CAG CTG TTC TGC CTG GCC CGG GCC TTC GTG AGG AAG ACC AGC Gln Arg Gln Leu Phe Cys Leu Ala Arg Ala Phe Val Arg Lys Thr Ser 1405 1410 1415	4275
ATC TTC ATC ATG GAT GAA GCA ACC GCC TCC ATC GAC ATG GCT ACG GAG Ile Phe Ile Met Asp Glu Ala Thr Ala Ser Ile Asp Met Ala Thr Glu 1420 1425 1430	4323
AAC ATC CTC CAG AAG GTG GTG ATG ACA GCC TTC GCA GAC CGC ACG GTG Asn Ile Leu Gln Lys Val Val Met Thr Ala Phe Ala Asp Arg Thr Val 1435 1440 1445	4371
GTC ACC ATC GCG CAT CGT GTG CAC ACC ATC CTG AGT GCA GAC CTG GTG Val Thr Ile Ala His Arg Val His Thr Ile Leu Ser Ala Asp Leu Val 1450 1455 1460 1465	4419
ATG GTC CTC AAG AGG GGT GCT ATC CTG GAG TTT GAC AAG CCA GAG ACG Met Val Leu Lys Arg Gly Ala Ile Leu Glu Phe Asp Lys Pro Glu Thr 1470 1475 1480	4467

SUBSTITUTE SHEET (RULE 26)

74

CTC CTC AGC CAG AAG GAC AGC GTG TTC GCC TCC TTT GTC CGT GCG GAC 4515  
 Leu Leu Ser Gln Lys Asp Ser Val Phe Ala Ser Phe Val Arg Ala Asp  
           1485                                  1490                                  1495

AAG TGA CTT ACC GAG CCA AAG T GCC ACC CCG C GCT CG CTT G CTT GC CTA GG 4568  
 Lys

ATT TCT AACT GCAA ATCA CT TGTA AATA AA TTA ATT CTT T GCT AAAAAA AAAAAA 4625

## (8) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1498 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met Pro Leu Ala Phe Cys Gly Thr Glu Asn His Ser Ala Ala Tyr Arg  
   1                          5                          10                          15  
 Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn  
                           20                          25                          30  
 Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe  
                           35                          40                          45  
 Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr  
                           50                          55                          60  
 Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Ile  
                           65                          70                          75                          80  
 Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp  
                           85                          90                          95  
 Gly Val Thr Glu Ser Arg His Leu His Leu Tyr Met Pro Ala Gly Met  
                           100                          105                          110  
 Ala Phe Met Ala Ala Ile Thr Ser Val Val Tyr Tyr His Asn Ile Glu  
                           115                          120                          125  
 Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr  
                           130                          135                          140  
 Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Tyr Asp His  
                           145                          150                          155                          160  
 Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val  
                           165                          170                          175  
 Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val  
                           180                          185                          190  
 Arg Arg Tyr Ile Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu  
                           195                          200                          205  
 Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu  
                           210                          215                          220  
 Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala  
                           225                          230                          235                          240

SUBSTITUTE SHEET (RULE 26)



His Lys Lys Pro Ile Asp Leu Arg Ala Ile Ala Lys Leu Pro Ile Ala  
 245 250 255  
 Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Val Ala Phe Asp Ala  
 260 265 270  
 Gln Ala Arg Lys Asp Thr Gln Ser Pro Gln Gly Ala Arg Ala Ile Trp  
 275 280 285  
 Arg Ala Leu Cys His Ala Phe Gly Arg Arg Leu Ile Leu Ser Ser Thr  
 290 295 300  
 Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile  
 305 310 315 320  
 Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn His Val Phe Gln Pro  
 325 330 335  
 Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu  
 340 345 350  
 Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu  
 355 360 365  
 Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly  
 370 375 380  
 Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met  
 385 390 395 400  
 His Met Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln  
 405 410 415  
 Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe  
 420 425 430  
 Phe Leu Cys Pro Asn Leu Trp Thr Met Pro Val Gln Ile Ile Val Gly  
 435 440 445  
 Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala  
 450 455 460  
 Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys  
 465 470 475 480  
 Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu His Ser Asn Glu Arg Leu  
 485 490 495  
 Lys Gln Thr Asn Glu Met Leu Arg Gly Met Lys Leu Leu Lys Leu Tyr  
 500 505 510  
 Ala Trp Glu Ser Ile Phe Cys Ser Arg Val Glu Val Thr Arg Arg Lys  
 515 520 525  
 Glu Met Thr Ser Leu Arg Ala Phe Ala Val Tyr Thr Ser Ile Ser Ile  
 530 535 540  
 Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val  
 545 550 555 560  
 Gly His Val Ser Phe Phe Lys Glu Ser Asp Leu Ser Pro Ser Val Ala  
 565 570 575  
 Phe Ala Ser Leu Ser Leu Phe His Ile Leu Val Thr Pro Leu Phe Leu  
 580 585 590  
 Leu Ser Ser Val Val Arg Ser Thr Val Lys Ala Leu Val Ser Val Gln

SUBSTITUTE SHEET (RULE 26)

595					600					605					
Lys	Leu	Ser	Glu	Phe	Leu	Ser	Ser	Ala	Glu	Ile	Arg	Glu	Glu	Gln	Cys
610					615					620					
Ala	Pro	Arg	Glu	Pro	Ala	Pro	Gln	Gly	Gln	Ala	Gly	Lys	Tyr	Gln	Ala
625					630					635					640
Val	Pro	Leu	Lys	Val	Val	Asn	Arg	Lys	Arg	Pro	Ala	Arg	Glu	Glu	Val
				645					650					655	
Arg	Asp	Leu	Leu	Gly	Pro	Leu	Gln	Arg	Leu	Ala	Pro	Ser	Met	Asp	Gly
			660					665					670		
Asp	Ala	Asp	Asn	Phe	Cys	Val	Gln	Ile	Ile	Gly	Gly	Phe	Phe	Thr	Trp
		675					680					685			
Thr	Pro	Asp	Gly	Ile	Pro	Thr	Leu	Ser	Asn	Ile	Thr	Ile	Arg	Ile	Pro
	690					695					700				
Arg	Gly	Gln	Leu	Thr	Met	Ile	Val	Gly	Gln	Val	Gly	Cys	Gly	Lys	Ser
705					710					715					720
Ser	Leu	Leu	Leu	Ala	Thr	Leu	Gly	Glu	Met	Gln	Lys	Val	Ser	Gly	Ala
				725					730					735	
Val	Phe	Trp	Asn	Ser	Asn	Leu	Pro	Asp	Ser	Glu	Gly	Arg	Gly	Pro	Gln
			740					745					750		
Gln	Pro	Arg	Ala	Gly	Asp	Ser	Ser	Trp	Leu	Gly	Tyr	Gln	Glu	Gln	Arg
		755					760					765			
Pro	Arg	Gly	Tyr	Ala	Ser	Gln	Lys	Pro	Trp	Leu	Leu	Asn	Ala	Thr	Val
	770					775					780				
Glu	Glu	Asn	Ile	Thr	Phe	Glu	Ser	Pro	Phe	Asn	Pro	Gln	Arg	Tyr	Lys
785					790					795					800
Met	Val	Ile	Glu	Ala	Cys	Ser	Leu	Gln	Pro	Asp	Ile	Asp	Ile	Leu	Pro
				805					810					815	
His	Gly	Asp	Gln	Thr	Gln	Ile	Gly	Glu	Arg	Gly	Ile	Asn	Leu	Ser	Gly
			820				825						830		
Gly	Gln	Arg	Pro	Asp	Gln	Cys	Gly	Pro	Glu	Pro	Ser	Thr	Ser	Arg	Pro
		835					840					845			
Met	Phe	Val	Phe	Leu	Asp	Asp	Pro	Phe	Ser	Ala	Leu	Asp	Val	His	Leu
	850					855					860				
Ser	Asp	His	Leu	Met	Gln	Ala	Gly	Ile	Leu	Glu	Leu	Leu	Arg	Asp	Asp
865					870					875					880
Lys	Arg	Thr	Val	Val	Leu	Val	Thr	His	Lys	Leu	Gln	Tyr	Leu	Pro	His
			885						890					895	
Ala	Asp	Trp	Ile	Ile	Ala	Met	Lys	Asp	Gly	Thr	Ile	Gln	Arg	Glu	Gly
			900					905					910		
Thr	Leu	Lys	Asp	Phe	Gln	Arg	Ser	Glu	Cys	Gln	Leu	Phe	Glu	His	Trp
		915					920					925			
Lys	Thr	Leu	Met	Asn	Arg	Gln	Asp	Gln	Glu	Leu	Glu	Lys	Glu	Thr	Val
	930					935					940				
Met	Glu	Arg	Lys	Ala	Ser	Glu	Pro	Ser	Gln	Gly	Leu	Pro	Arg	Ala	Met
945						950					955				960

SUBSTITUTE SHEET (RULE 26)

Ser Ser Arg Asp Gly Leu Leu Leu Asp Glu Glu Glu Glu Glu Glu  
 965 970 975  
 Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Val Leu His Gln  
 980 985 990  
 Arg Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala  
 995 1000 1005  
 Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His  
 1010 1015 1020  
 Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser  
 1025 1030 1035 1040  
 Ala Leu Val Leu Ser Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu  
 1045 1050 1055  
 Cys Asp Leu Asp Gln Ser Val Tyr Ala Met Val Phe Thr Leu Leu Cys  
 1060 1065 1070  
 Ser Leu Gly Ile Val Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp  
 1075 1080 1085  
 Thr Gly Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg  
 1090 1095 1100  
 Ile Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser  
 1105 1110 1115 1120  
 Ile Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile  
 1125 1130 1135  
 Pro Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser  
 1140 1145 1150  
 Ala Leu Thr Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu  
 1155 1160 1165  
 Leu Pro Leu Ala Val Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val  
 1170 1175 1180  
 Ala Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu  
 1185 1190 1195 1200  
 Val Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala  
 1205 1210 1215  
 Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp  
 1220 1225 1230  
 Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu  
 1235 1240 1245  
 Glu Val Cys Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala  
 1250 1255 1260  
 Ala Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu  
 1265 1270 1275 1280  
 Val Gly Leu Gly Leu Thr Tyr Ala Leu Met Ile Gly Ile Cys Gly Arg  
 1285 1290 1295  
 Thr Ala Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe Phe Arg Met Val  
 1300 1305 1310  
 Asp Met Phe Glu Gly Arg Ile Ile Ile Asp Gly Ile Asp Ile Ala Lys

SUBSTITUTE SHEET (RULE 26)

1315	1320	1325
Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile Ile Leu Gln Asp		
1330	1335	1340
Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu Asp Pro Glu Lys		
1345	1350	1355 1360
Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu Ile Ala Gln Leu		
1365	1370	1375
Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp Ala Ile Ile Thr		
1380	1385	1390
Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln Leu Phe Cys Leu		
1395	1400	1405
Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile Met Asp Glu Ala		
1410	1415	1420
Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu Gln Lys Val Val		
1425	1430	1435 1440
Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile Ala His Arg Val		
1445	1450	1455
His Thr Ile Leu Ser Ala Asp Leu Val Met Val Leu Lys Arg Gly Ala		
1460	1465	1470
Ile Leu Glu Phe Asp Lys Pro Glu Thr Leu Leu Ser Gln Lys Asp Ser		
1475	1480	1485
Val Phe Ala Ser Phe Val Arg Ala Asp Lys		
1490	1495	

## (9) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 8 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Pro Leu Ala Phe Cys Gly Thr  
 1 5

## (10) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 11 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Asn His Ser Ala Ala Tyr Arg Val Asp Gln Gly  
 1 5 10

## (11) INFORMATION FOR SEQ ID NO:10:

SUBSTITUTE SHEET (RULE 26)

79

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 26 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAGAGAAGCT TNTGNGGNGA NAANCA

26

- (12) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 28 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GAGAGAGAAT TCCNTGNTCN ACNCNNTA

28

- (13) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 47 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTTTGCGGGA CGGAGAATCA CTCGGCCGCC TACCGCGTCG ACCAAGG

47

- (14) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 9 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: cDNA

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GCCNCCAUG

9

- (15) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 7 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GXXGXGK

SUBSTITUTE SHEET (RULE 26)

80

## (16) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 4 amino acids
  - (B) TYPE: amino acid
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

XXXD

## (17) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acids

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CACGCTCAGG TTCTGGAT

18

## (18) INFORMATION FOR SEQ ID NO:17:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acids

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

TCAACTGGAT GGTGAGGA

18

## (17) INFORMATION FOR SEQ ID NO:18:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

SUBSTITUTE SHEET (RULE 26)

81

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGACATCGCC AACTGC

17

(18) INFORMATION FOR SEQ ID NO:19:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TCCTGGCAGT GCCTTCA

17

(19) INFORMATION FOR SEQ ID NO:20:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 20 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TCCTCTCAGG GTCCAGGTTA

20

(20) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 17 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

ACAAGGAGCC TGGGGAT

17

(21) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:

SUBSTITUTE SHEET (RULE 26)

82

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TGCATGGGTC CCAAGTA

17

(22) INFORMATION FOR SEQ ID NO:23:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 25 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TTGACCATTTC ACCACATTGG TGTGC

25

(23) INFORMATION FOR SEQ ID NO:24:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 17 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: YES

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

TCCTGGCAGT GCCTTCA

17

(24) INFORMATION FOR SEQ ID NO:25:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 18 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

SUBSTITUTE SHEET (RULE 26)



(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

ACCATCGACC AGCACATC

18

(25) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1308 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GGACCTGCAG CAGCTGGATG ACACCACCCA GCTTCCACTT CTCTCACACT TTGCCGAAAC	60
CGTAGAAGGA CTCACCACCA TCCGGGCCTT CAGGTATGAG GCCCGGTTCC AGCAGAAGCT	120
TCTCGAATAC ACAGACTCCA ACAACATTGC TTCCCTCTTC CTCACAGCTG CCAACAGATG	180
GCTGGAAGTC CGAATGGAGT ACATCGGTGC ATGTGTGGTG CTCATCGCAG CGGTGACCTC	240
CATCTCCAAC TCCCTGCACA GGGAGCTCTC TGCTGGCCTG GTGGGCCTGG GCCTTACCTA	300
CGCCCTAATG GTCTCCAAC TACCTCAACTG GATGGTGAGG AACCTGGCAG ACATGGAGCT	360
CCAGCTGGGG GCTGTGAAGC GCATCCATGG GCTCCTGAAA ACCGAGGCAG AGAGCTACGA	420
GGGACTCCTG GCACCATCGC TGATCCCAA GAACTGGCCA GACCAAGGGA AGATCCAGAT	480
CCAGAACCTG AGCGTGCGCT ACGACAGCTC CCTGAAGCCG GTGCTGAAGC ACGTCAATGC	540
CCTCATCTCC CCTGGACAGA AGATCGGGAT CTGCGGCCGC ACCGGCAGTG GGAAGTCCTC	600
CTTCTCTCTT GCCTTCTTCC GCATGGTGGA CACGTTCGAA GGGCACATCA TCATTGATGG	660
CATTGACATC GCCAAACTGC CGCTGCACAC CCTGCGCTCA CGCCTCTCCA TCATCCTGCA	720
GGACCCCGTC CTCTTCAGCG GCACCATCCG ATTTAACCTG GACCCTGAGA GGAAGTGCTC	780
AGATAGCACA CTGTGGGAGG CCCTGGAAAT CGCCCAGCTG AAGCTGGTGG TGAAGGCACT	840
GCCAGGAGGC CTCGATGCCA TCATCACAGA AGGCGGGGAG AATTTCAGCC AGGGACAGAG	900
GCAGCTGTTC TGCCTGGCCC GGGCCTTCGT GAGGAAGACC AGCATCTTCA TCATGGACGA	960
GGCCACGGCT TCCATTGACA TGGCCACGGA AAACATCCTC CAAAAGGTGG TGATGACAGC	1020
CTTCGCAGAC CGCACTGTGG TCACCATCGC GCATCGAGTG CACACCATCC TGAGTGCAGA	1080
CCTGGTGATC GTCCTGAAGC GGGGTGCCAT CCTTGAGTTC GATAAGCCAG AGAAGCTGCT	1140

SUBSTITUTE SHEET (RULE 26)

CAGCCGGAAG GACAGCGTCT TCGCCTCCTT CGTCCGTGCA GACAAGTGAC CTGCCAGAGC 1200  
 CCAAGTGCCA TCCCACATTC GGACCTGCC CATACCCCTG CCTGGGTTTT CTAAGTGTA 1260  
 ATCACTTGTA AATAAATAGA TTTGATTATT TCCTAAAAAA AAAAAAAA 1308

## (26) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1308 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 2..1186

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

G GAC CTG CAG CAG CTG GAT GAC ACC ACC CAG CTT CCA CTT CTC TCA	46
Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu Leu Ser	
1 5 10 15	
CAC TTT GCC GAA ACC GTA GAA GGA CTC ACC ACC ATC CGG GCC TTC AGG	94
His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala Phe Arg	
20 25 30	
TAT GAG GCC CGG TTC CAG CAG AAG CTT CTC GAA TAC ACA GAC TCC AAC	142
Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp Ser Asn	
35 40 45	
AAC ATT GCT TCC CTC TTC CTC ACA GCT GCC AAC AGA TGG CTG GAA GTC	190
Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu Glu Val	
50 55 60	
CGA ATG GAG TAC ATC GGT GCA TGT GTG GTG CTC ATC GCA GCG GTG ACC	238
Arg Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala Val Thr	
65 70 75	
TCC ATC TCC AAC TCC CTG CAC AGG GAG CTC TCT GCT GGC CTG GTG GGC	286
Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu Val Gly	
80 85 90 95	
CTG GGC CTT ACC TAC GCC CTA ATG GTC TCC AAC TAC CTC AAC TGG ATG	334
Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn Trp Met	
100 105 110	
GTG AGG AAC CTG GCA GAC ATG GAG CTC CAG CTG GGG GCT GTG AAG CGC	382
Val Arg Asn Leu Ala Asp Met Glu Leu Gln Leu Gly Ala Val Lys Arg	
115 120 125	
ATC CAT GGG CTC CTG AAA ACC GAG GCA GAG AGC TAC GAG GGA CTC CTG	430
Ile His Gly Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly Leu Leu	
130 135 140	
GCA CCA TCG CTG ATC CCA AAG AAC TGG CCA GAC CAA GGG AAG ATC CAG	478
Ala Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys Ile Gln	
145 150 155	

SUBSTITUTE SHEET (RULE 26)

85

ATC Ile 160	CAG Gln 160	AAC Asn 160	CTG Leu 160	AGC Ser 165	GTG Val 165	CGC Arg 165	TAC Tyr 165	GAC Asp 170	AGC Ser 170	TCC Ser 170	CTG Leu 170	AAG Lys 170	CCG Pro 170	GTG Val 175	CTG Leu 175	526
AAG Lys 180	CAC His 180	GTC Val 180	AAT Asn 180	GCC Ala 180	CTC Leu 180	ATC Ile 180	TCC Ser 185	CCT Pro 185	GGA Gly 185	CAG Gln 185	AAG Lys 190	ATC Ile 190	GGG Gly 190	ATC Ile 190	TGC Cys 190	574
GGC Gly 195	CGC Arg 195	ACC Thr 195	GGC Gly 195	AGT Ser 195	GGG Gly 195	AAG Lys 200	TCC Ser 200	TCC Ser 200	TTC Phe 200	TCT Ser 205	CTT Leu 205	GCC Ala 205	TTC Phe 205	TTC Phe 205	CGC Arg 205	622
ATG Met 210	GTG Val 210	GAC Asp 210	ACG Thr 210	TTC Phe 215	GAA Glu 215	GGG Gly 215	CAC His 215	ATC Ile 215	ATC Ile 220	ATT Ile 220	GAT Asp 220	GGC Gly 220	ATT Ile 220	GAC Asp 220	ATC Ile 220	670
GCC Ala 225	AAA Lys 225	CTG Leu 225	CCG Pro 230	CTG Leu 230	CAC His 230	ACC Thr 230	CTG Leu 235	CGC Arg 235	TCA Ser 235	CGC Arg 235	CTC Leu 235	TCC Ser 235	ATC Ile 235	ATC Ile 235	CTG Leu 235	718
CAG Gln 240	GAC Asp 240	CCC Pro 240	GTC Val 245	CTC Leu 245	TTC Phe 245	AGC Ser 245	GGC Gly 250	ACC Thr 250	ATC Ile 250	CGA Arg 250	TTT Phe 250	AAC Asn 255	CTG Leu 255	GAC Asp 255	CCT Pro 255	766
GAG Glu 260	AGG Arg 260	AAG Lys 260	TGC Cys 260	TCA Ser 260	GAT Asp 265	AGC Ser 265	ACA Thr 265	CTG Leu 265	TGG Trp 265	GAG Glu 270	GCC Ala 270	CTG Leu 270	GAA Glu 270	ATC Ile 270	GCC Ala 270	814
CAG Gln 275	CTG Leu 275	AAG Lys 275	CTG Leu 275	GTG Val 280	GTG Val 280	AAG Lys 280	GCA Ala 280	CTG Leu 280	CCA Pro 285	GGA Gly 285	GGC Gly 285	CTC Leu 285	GAT Asp 285	GCC Ala 285	ATC Ile 285	862
ATC Ile 290	ACA Thr 290	GAA Glu 290	GGC Gly 295	GGG Gly 295	GAG Glu 295	AAT Asn 295	TTC Phe 295	AGC Ser 295	CAG Gln 300	GGA Gly 300	CAG Gln 300	AGG Arg 300	CAG Gln 300	CTG Leu 300	TTC Phe 300	910
TGC Cys 305	CTG Leu 305	GCC Ala 310	CGG Arg 310	GCC Ala 310	TTC Phe 310	GTG Val 310	AGG Arg 315	AAG Lys 315	ACC Thr 315	AGC Ser 315	ATC Ile 315	TTC Phe 315	ATC Ile 315	ATG Met 315	GAC Asp 315	958
GAG Glu 320	GCC Ala 320	ACG Thr 325	GCT Ala 325	TCC Ser 325	ATT Ile 325	GAC Asp 330	ATG Met 330	GCC Ala 330	ACG Thr 330	GAA Glu 330	AAC Asn 335	ATC Ile 335	CTC Leu 335	CAA Gln 335	AAG Lys 335	1006
GTG Val 340	GTG Val 340	ATG Met 340	ACA Thr 340	GCC Ala 340	TTC Phe 340	GCA Ala 345	GAC Asp 345	CGC Arg 345	ACT Thr 345	GTG Val 350	GTC Val 350	ACC Thr 350	ATC Ile 350	GCG Ala 350	CAT His 350	1054
CGA Arg 355	GTG Val 355	CAC His 355	ACC Thr 355	ATC Ile 360	CTG Leu 360	AGT Ser 360	GCA Ala 360	GAC Asp 360	CTG Leu 365	GTG Val 365	ATC Ile 365	GTC Val 365	CTG Leu 365	AAG Lys 365	CGG Arg 365	1102
GGT Gly 370	GCC Ala 370	ATC Ile 370	CTT Leu 375	GAG Glu 375	TTC Phe 375	GAT Asp 375	AAG Lys 375	CCA Pro 380	GAG Glu 380	AAG Lys 380	CTG Leu 380	CTC Leu 380	AGC Ser 380	CGG Arg 380	AAG Lys 380	1150
GAC Asp 385	AGC Ser 385	GTC Val 390	TTC Phe 390	GCC Ala 390	TCC Ser 390	TTC Phe 390	GTC Val 390	CGT Arg 395	GCA Ala 395	GAC Asp 395	AAG Lys 395	TGACCTGCCA				1196
GAGCCCAAGT GCCATCCCAC ATTCGGACCC TGCCCATACC CCTGCCTGGG TTTTCTAACT																1256
GTAAATCACT TGTAATAAAA TAGATTTGAT TATTTCTAA AAAAAAAAAA AA																1308

(27) INFORMATION FOR SEQ ID NO:28:

SUBSTITUTE SHEET (RULE 26)

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 395 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu Leu Ser His  
 1 5 10 15  
 Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala Phe Arg Tyr  
 20 25 30  
 Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp Ser Asn Asn  
 35 40 45  
 Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu Glu Val Arg  
 50 55 60  
 Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala Val Thr Ser  
 65 70 75 80  
 Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu Val Gly Leu  
 85 90 95  
 Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn Trp Met Val  
 100 105 110  
 Arg Asn Leu Ala Asp Met Glu Leu Gln Leu Gly Ala Val Lys Arg Ile  
 115 120 125  
 His Gly Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly Leu Leu Ala  
 130 135 140  
 Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys Ile Gln Ile  
 145 150 155 160  
 Gln Asn Leu Ser Val Arg Tyr Asp Ser Ser Leu Lys Pro Val Leu Lys  
 165 170 175  
 His Val Asn Ala Leu Ile Ser Pro Gly Gln Lys Ile Gly Ile Cys Gly  
 180 185 190  
 Arg Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe Phe Arg Met  
 195 200 205  
 Val Asp Thr Phe Glu Gly His Ile Ile Ile Asp Gly Ile Asp Ile Ala  
 210 215 220  
 Lys Leu Pro Leu His Thr Leu Arg Ser Arg Leu Ser Ile Ile Leu Gln  
 225 230 235 240  
 Asp Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu Asp Pro Glu  
 245 250 255  
 Arg Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu Ile Ala Gln  
 260 265 270  
 Leu Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp Ala Ile Ile  
 275 280 285  
 Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln Leu Phe Cys  
 290 295 300  
 Leu Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile Met Asp Glu

SUBSTITUTE SHEET (RULE 26)

305				310				315				320			
Ala	Thr	Ala	Ser	Ile	Asp	Met	Ala	Thr	Glu	Asn	Ile	Leu	Gln	Lys	Val
				325					330					335	
Val	Met	Thr	Ala	Phe	Ala	Asp	Arg	Thr	Val	Val	Thr	Ile	Ala	His	Arg
				340					345					350	
Val	His	Thr	Ile	Leu	Ser	Ala	Asp	Leu	Val	Ile	Val	Leu	Lys	Arg	Gly
				355					360					365	
Ala	Ile	Leu	Glu	Phe	Asp	Lys	Pro	Glu	Lys	Leu	Leu	Ser	Arg	Lys	Asp
				370					375					380	
Ser	Val	Phe	Ala	Ser	Phe	Val	Arg	Ala	Asp	Lys					
				385					390					395	

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 195 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(iv) ANTI-SENSE: NO

CCATGCCTGG TGGCTGAGCC CAGCCCAGCC CCCAGACCA TCCTGATCC CAAAGAACTG	60
GCCAGACCAA GGGAAGATCC AGATCCAGAA CCTGAGCGTG CGCTACGACA GCTCCCTGAA	120
GCCGGTGCTG AAGCAGCTCA ATGCCCTCAT CTCCCCTGGA CAGAAGGTCA GTGCACGGGC	180
CCAACCCAAT GCTGC	195

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2454 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iv) ANTI-SENSE: NO

CTGCTCTCTG	CAGCCAGACA	TCGACATCCT	GCCCCATGGA	GACCAGACCC	AGATTGGGGA	60
ACGGGGCATC	AACCTGTCTG	GTGGTCAACG	CCAGCGAATC	AGTGTGGCCC	GAGCCCTCTA	120
CCAGCACGCC	AACGTTGTCT	TCTTGGAATG	CCCCTTCTCA	GCTCTGGATA	TCCATCTGAG	180

BNSDOCID: <WO\_\_9528411A1\_I\_>

TGACCACTTA	ATGCAGGCCG	GCATCCTTGA	GCTGCTCCGG	GACGACAAGA	GGACAGTGGT	240
CTTAGTGACC	CACAAGCTAC	AGTACCTGCC	CCATGCAGAC	TGGATCATTG	CCATGAAGGA	300
TGGCACCATC	CAGAGGGAGG	GTACCCTCAA	GGACTTCCAG	AGGTCTGAAT	GCCAGCTCTT	360
TGAGCACTGG	AAGACCCTCA	TGAACCGACA	GGACCAAGAG	CTGGAGAAGG	AGACTGTCAC	420
AGAGAGAAAA	GCCACAGAGC	CACCCAGGG	CCTATCTCGT	GCCATGTCCT	CGAGGGATGG	480
CCTTCTGCAG	GATGAGGAAG	AGGAGGAAGA	GGAGGCAGCT	GAGAGCGAGG	AGGATGACAA	540
CCTGTGCTCC	ATGCTGCACC	AGCGTGCTGA	GATCCCATGG	CGAGCCTGCG	CCAAGTACCT	600
GTCCTCCGCC	GGCATCCTGC	TCTGTCTGTT	GCTGGTCTTC	TCACAGCTGC	TCAAGCACAT	660
GGTCCTGGTG	GCCATCGACT	ACTGGCTGGC	CAAGTGGACC	GACAGCGCCC	TGACCCTGAC	720
CCCTGCAGCC	AGGAACTGCT	CCCTCAGCCA	GGAGTGCACC	CTCGACCAGA	CTGTCTATGC	780
CATGGTGTTC	ACGGTGCTCT	GCAGCCTGGG	CATTGTGCTG	TGCCTCGTCA	CGTCTGTCAC	840
TGTGGAGTGG	ACAGGGCTGA	AGGTGGCCAA	GAGACTGCAC	CGCAGCCTGC	TAAACCGGAT	900
CATCCTAGCC	CCCATGAGGT	TTTTTGAGAC	CACTCCCCTT	GGGAGCATCC	TGAACAGATT	960
TTCATCTGAC	TGTAACACCA	TCGACCAGCA	CATCCCATCC	ACGCTGGAGT	GCCTGAGCCG	1020
CTCCACCCTG	CTCTGTGTCT	CAGCCCTGGC	CGTCATCTCC	TATGTCACAC	CTGTGTTCTT	1080
CGTGGCCCTC	CTTCCCCTGG	CCATCGTGTG	CTACTTCATC	CAGAACTACT	TCCGGGTGGC	1140
GTCCAGGGAC	CTGCAGCAGC	TGGATGACAC	CACCCAGCTT	CCACTTCTCT	CACACTTTGC	1200
CGAAACCGTA	GAAGGACTCA	CCACCATCCG	GGCCTTCAGG	TATGAGGCC	GGTTCCAGCA	1260
GAAGCTTCTC	GAATACACAG	ACTCCAACAA	CATTGCTTCC	CTCTTCCTCA	CAGCTGCCAA	1320
CAGATGGCTG	GAAGTCCGAA	TGGAGTACAT	CGGTGCATGT	GTGGTGCTCA	TCGCAGCGGT	1380
GACCTCCATC	TCCAACCTCC	TGCACAGGGA	GCTCTCTGCT	GGCCTGGTGG	GCCTGGGCC	1440
TACCTACGCC	CTAATGGTCT	CCAACTACCT	CAACTGGATG	GTGAGGAACC	TGGCAGACAT	1500
GGAGCTCCAG	CTGGGGGCTG	TGAAGCGCAT	CCATGGGCTC	CTGAAAACCG	AGGCAGAGAG	1560
CTACGAGGGA	CTCCTGGCAC	CATCGCTGAT	CCCAAAGAAC	TGGCCAGACC	AAGGGAAGAT	1620
CCAGATCCAG	AACCTGAGCG	TGCGCTACGA	CAGCTCCCTG	AAGCCGGTGC	TGAAGCACGT	1680
CAATGCCCTC	ATCTCCCCTG	GACAGAAGAT	CGGGATCTGC	GGCCGCACCG	GCAGTGGGAA	1740
GTCCTCCTTC	TCTCTTGCTT	TCTTCCGCAT	GGTGGACACG	TTCGAAGGGC	ACATCATCAT	1800
TGATGGCATT	GACATCGCCA	AACTGCCGCT	GCACACCCTG	CGCTCACGCC	TCTCCATCAT	1860
CCTGCAGGAC	CCCGTCTCTT	TCAGCGGCAC	CATCCGATTT	AACCTGGACC	CTGAGAGGAA	1920
GTGCTCAGAT	AGCACACTGT	GGGAGGCCCT	GGAAATCGCC	CAGCTGAAGC	TGGTGGTGAA	1980
GGCACTGCCA	GGAGGCCTCG	ATGCCATCAT	CACAGAAGGC	GGGGAGAATT	TCAGCCAGGG	2040
ACAGAGGCAG	CTGTTCTGCC	TGGCCCGGGC	CTTCGTGAGG	AAGACCAGCA	TCTTCATCAT	2100
GGACGAGGCC	ACGGCTTCCA	TTGACATGGC	CACGGAAAAC	ATCCTCCAAA	AGGTGGTGAT	2160
GACAGCCTTC	GCAGACCGCA	CTGTGGTCAC	CATCGCGCAT	CGAGTGCACA	CCATCCTGAG	2220

SUBSTITUTE SHEET (RULE 26)

\_\_\_\_\_

|

\_\_\_\_\_

|

TGCAGACCTG GTGATCGTCC TGAAGCGGGG TGCCATCCTT GAGTTCGATA AGCCAGAGAA 2280  
 GCTGCTCAGC CGGAAGGACA GCGTCTTCGC CTCCTTCGTC CGTGCAGACA AGTGACCTGC 2340  
 CAGAGCCCAA GTGCCATCCC ACATTCGGAC CCTGCCATA CCCCTGCCTG GGTTCCTAA 2400  
 CTGTAAATCA CTTGTAAATA AATAGATTG ATTATTCCT AAAAAAAAAA AAAA 2454

## (30) INFORMATION FOR SEQ ID NO:31:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2294 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GATGACCCCT TCTCAGCTTT GGATGTCCAT CTGAGTGACC ACCTGATGCA GGCCGGCATC 60  
 CTTGAGCTGC TCCGGGATGA CAAGAGGACA GTGGTCTTGG TGACCCACAA GCTACAGTAT 120  
 CTGCCTCATG CAGACTGGAT CATTGCCATG AAGGATGGGA CCATTCAGAG GGAAGGGACG 180  
 CTCAAGGACT TCCAGAGGTC CGAGTGCCAG CTCTTTGAGC ACTGGAAGAC CCTCATGAAC 240  
 CGGCAGGACC AAGAGCTGGA GAAGGAGACA GTCATGGAGA GGAAAGCCTC AGAGCCATCT 300  
 CAGGGCCTGC CCCGTGCCAT GTCCTCCAGA GACGGCCTTC TGCTGGATGA GGAAGAGGAG 360  
 GAAGAGGAGG CAGCCGAAAG CGAGGAAGAT GACAACCTAT CTTCACTGCT GCATCAGCGA 420  
 GCTAAGATCC CCTGGCGAGC CTGCACTAAG TATCTGTCCT CTGTGGCAT TCTGCTCCTG 480  
 TCCCTGCTTG TCTTCTCCCA GCTGCTCAAG CACATGGTCT TGGTGGCCAT TGATTATTGG 540  
 CTGGCCAACT GGACGACAG TGCCCTGGTC CTGAGCCCCG CTGCCAGGAA CTGTTTCGCTC 600  
 AGCCAGGAAT GTGACCTGGA CCAGTCTGTC TATGCCATGG TATTCACCTT GCTCTGCAGC 660  
 CTGGGTATCG TGCTGTGCCT GGTACCTCT GTCAGTGTGG AGTGGACGGG ACTGAAGGTG 720  
 GCCAAGAGGC TACACCGCAG CCTGCTCAAC CGCATCATCC TGGCCCCCAT GAGGTTCTTT 780  
 GAGACCACAC CCCTCGGGAG TATCCTGAAC AGATTTTCAT CCGACTGTAA CACCATTGAC 840  
 CAGCACATCC CATCCACGCT GGAGTGTCTG AGCCGGTCCA CCCTGCTGTG TGTCTCCGCC 900  
 CTGACTGTCA TCTCCTATGT CACACCCGTG TTCTCGTGG CCCTCTTACC CCTAGCTGTT 960  
 GTGTGCTACT TCATTAGAA GTACTTCCGA GTGGCATCCA GGGACCTGCA GCAGCTGGAC 1020  
 GACACGACGC AGCTCCCGCT CGTCTCACAC TTTGCTGAAA CTGTGGAGGG ACTCACCACC 1080  
 ATCCGTGCCT TCAGGTACGA GGCCCGGTTT CAGCAGAAGC TTCTAGAATA TACCGACTCC 1140  
 AACAAATCG CCTCCCTCTT CCTCAGGCA GCCAACAGAT GGCTGGAAGT CTGCATGGAG 1200  
 TACATCGGAG CGTGCGTGGT ACTCATTGCG GTTGGCAGCT CCATCTCCAA CTCCCTGCAC 1260

SUBSTITUTE SHEET (RULE 26)



```

AGGGAAC TTT CTGCTGGCCT GGTGGGCCTG GGCCTCACCT ATGCCTTGAT GGTCTCCAAC 1320
TACCTCAACT GGATGGTGAG GAACCTGGCG GACATGGAGA TCCAGCTGGG GGCTGTGAAG 1380
AGGATCCACG CACTCCTGAA AACCAGGGCG GAGAGCTATG AGGGGCTCCT GGCGCCGTCG 1440
TTGATCCCCA AGAACTGGCC AGACCAAGGG AAGATCCAAA TTCAGAACCT GAGCGTGCGC 1500
TATGACAGCT CCCTGAAGCC AGTGCTGAAG CATGTCAACA CCCTCATCTC CCCGGGGCAG 1560
AAGATCGGGA TCTGCGGCCG CACAGGCAGC GGAAGTCCT CCTTCTCCCT GGCTTTTTTC 1620
CGAATGGTGG ACATGTTTGA AGGACGCATC ATCATTGATG GCATCGACAT CGCCAAGCTG 1680
CCACTTCACA CGCTGCGCTC ACGCCTGTCC ATCATCCTAC AGGACCCCGT CCTCTTCAGC 1740
GGCACGATCA GATTCAACCT GGACCCCGAG AAGAAATGCT CAGACAGCAC ACTGTGGGAG 1800
GCCCTGGAGA TCGCCAGCT GAAGCTGGTA GTGAAGGCAC TGCCAGGAGG CCTAGATGCC 1860
ATCATCACAG AAGGAGGGGA GAATTTTAGC CAGGGCCAGA GGCAGCTGTT CTGCCTGGCC 1920
CGGGCCTTCG TGAGGAAGAC CAGCATCTTC ATCATGGATG AAGCAACCGC CTCCATCGAC 1980
ATGGCTACGG AGAACATCCT CCAGAAGGTG GTGATGACAG CCTTCGCAGA CCGCACGGTG 2040
GTCACCATCG CGCATCGTGT GCACACCATC CTGAGTGCAG ACCTGGTGAT GGTCTCAAG 2100
AGGGGTGCTA TCCTGGAGTT TGACAAGCCA GAGACGCTCC TCAGCCAGAA GGACAGCGTG 2160
TTCGCTCCTT TTGTCCGTGC GGACAAGTGA CTTACGGGAG CCAAAGTGCC ACCCCGCGCC 2220
TCGCTTGCTT GCCTAGGATT TCTAAGTGA AATCACTTGT AAATAAATTA ATTCTTTGCT 2280
AAAAAAAAA AAAA 2294

```

## (31) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 5110 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

```

CCCTTGTGAC AGGTCAGTCT TACGAGAATA TGGTAACTGA GATCATGTCA ATGGGCTATG 60
AACGAGAACA AGTAATTGCA GCCCTGAGAG CCAGCTTCAA CAACCCTGAT AGAGCTGTGG 120
AATATCTTCT AATGGGAATC CCTGGAGACT GAGGAGTTCC AGTACTCACA GCCTGTGGAG 180
GAGGATCAAC CACGGCCTGA CTTTCGCGGC CGCCGCGGGA GCGCGCGGA GCCGGAGCCG 240
AGCCCGTGCG CGCGCCACCA TGCCTTTGGC CTTCTGCGGC ACCGAGAACC ACTCGGCCGC 300
CTACCGGGTG GACCAAGGCG TCCTCAACAA CGGCTGCTTC GTGGACGCGC TCAATGTGGT 360
GCCACATGTC TTTCTGCTCT TCATCACCTT CCCCATCCTC TTCATCGGAT GGGGCAGCCA 420

```

SUBSTITUTE SHEET (RULE 26)

GAGCTCCAAG	GTGCACATTC	ACCACAGCAC	CTGGCTCCAT	TTCCCGGGGC	ACAACCTGCG	480
CTGGATCCTG	ACCTTCATAC	TGCTCTTCGT	CCTCGTGTGT	GAGATCGCTG	AGGGTATCCT	540
GTCTGACGGG	GTGACAGAAT	CCCGCCACCT	CCACTTATAC	ATGCCAGCTG	GGATGGCATT	600
CATGGCTGCC	ATCACCTCTG	TGGTCTACTA	CCATAACATT	GAGACCTCTA	ACTTTCCCAA	660
GCTGCTGATT	GCTCTGCTCA	TCTACTGGAC	CCTGGCCTTC	ATCACGAAGA	CCATCAAGTT	720
CGTCAAGTTC	TACGACCACG	CCATTGGCTT	CTCTCAGCTG	CGCTTCTGCC	TCACGGGGCT	780
TCTGGTGATC	CTCTACGGGA	TGCTGCTGCT	TGTGGAGGTC	AATGTGATCC	GGGTGAGGAG	840
ATACGTCTTC	TTCAAGACAC	CAAGGGAAGT	AAAGCCCCCC	GAGGACCTAC	AGGACCTGGG	900
TGTGCGCTTT	CTGCAGCCCT	TCGTAAACCT	GCTATCAAAG	GGGACCTACT	GGTGGATGAA	960
TGCCTTCATC	AAGACTGCTC	ACAAGAAGCC	CATCGACCTG	CGGGCCATCG	GGAAGCTGCC	1020
CATTGCCATG	AGAGCCCTCA	CCAACCTACCA	GCGACTCTGC	TTGGCCTTCG	ATGCCCAGGC	1080
GCGGAAGGAC	ACACAGAGCC	AGCAGGGTGC	CCGGGCCATC	TGGAGGGGTC	TCTGTCTATG	1140
CTTTGGGAGA	CGGCTGGTCC	TCAGCAGCAC	ATTCCGTATC	CTGGCCGACC	TCCTGGGCTT	1200
TGCTGGGCCA	CTCTGCATCT	TCGGGATCGT	GGACCACCTC	GGGAAGGAGA	ACCACGTCTT	1260
CCAGCCCCAAG	ACACAGTTTC	TTGGAGTTTA	CTTTGTCTCA	TCCAAGAGT	TCCTCGGCAA	1320
TGCCTATGTC	TTGGCTGTTC	TTCTGTTTCT	TGCCCTCCTG	CTGCAAAGGA	CCTTTCTACA	1380
AGCCTCGTAC	TACGTTGCCA	TTGAACTGG	GATCAACCTG	AGAGGAGCAA	TCCAGACCAA	1440
GATTTACAAT	AAGATCATGC	ACTTGTCTAC	TTCCAACCTG	TCCATGGGGG	AAATGACTGC	1500
TGGGCAGATC	TGCAACCTGG	TGGCCATCGA	CACCAACCAG	CTCATGTGGT	TTTTCTTCTT	1560
ATGCCCCAAC	CTCTGGGCTA	TGCCGGTACA	GATCATTGTG	GGCGTGATCC	TCCTCTACTA	1620
CATCCTTGGG	GTCAGCGCCT	TGATTGGAGC	GGCTGTCTATC	ATTCTGCTGG	CTCCTGTACA	1680
GTACTTTGTG	GCCACCAAGC	TGTCCCAGGC	ACAGCGGACG	ACCCTGGAAT	ATTCCAATGA	1740
GAGGCTGAAG	CAGACCAATG	AGATGCTCCG	GGGCATCAAG	TTGCTCAAGC	TCTATGCGTG	1800
GGAGAACATC	TTCTGCTCCA	GGGTGGAGAA	GACACGCAGG	AAGGAAATGA	CCAGCCTCAG	1860
GGCCTTCGCT	GTCTACACCT	CCATCTCCAT	CTTCATGAAC	ACAGCTATCC	CCATCGCTGC	1920
TGTCCTCATC	ACCTTCGTGG	GCCACGTCAG	CTTCTTCAAA	GAGTCGGACT	TCTCGCCCTC	1980
GGTGGCCTTT	GCCTCTCTCT	CTCTCTTCCA	CATCCTGGTC	ACACCGCTGT	TCCTGCTGTC	2040
TAGTGTGGTT	CGGTCCACTG	TCAAGGCCCT	GGTGAGCGTG	CAAAAGCTGA	GTGAGTTCCT	2100
GTCCAGTGCA	GAGATCCGTG	AGGAACAGTG	TGCCCCCGCA	GAGCCCGCAC	CCCAAGGCCA	2160
AGCGGGCAAG	TACCAGGCGG	TGCCCTCAA	GGTCGTAAAC	CGCAAGCGCC	CAGCCCGAGA	2220
AGAAGTCCGG	GACCTCTTGG	GCCCACTGCA	GAGGCTGACT	CCCAGCACGG	ATGGAGACGC	2280
TGACAACTTC	TGTGTCCAGA	TCATCGGAGG	CTTCTTCACC	TGGACCCCTG	ATGGAATCCC	2340
CACCCTGTCC	AACATCACCA	TCCGTATCCC	CCGAGGTCAG	CTGACCATGA	TCGTGGGGCA	2400
GGTGGGCTGT	GGCAAGTCCT	CGCTCCTTCT	GGCCACCCTG	GGGGAGATGC	AGAAGGTCTC	2460

SUBSTITUTE SHEET (RULE 26)

TGGAGCTGTC	TTCTGGAACA	GCCTTCCAGA	CAGCGAGGGG	AGAAGACCCC	AGCAACCCAG	2520
AGCGGGAGAC	AGCGGCCGAT	TCGGATGCCA	GGAGCAGAGG	CCCTGTGGCT	ACGCATCTCA	2580
GAAACCATGG	CTGCTAAATG	CCACTGTGGA	GGAGAACATC	ACCTTCGAGA	GTCCCTTCAA	2640
TAAGCAACGG	TACAAGATGG	TCATCGAAGC	CTGCTCCCTG	CAGCCAGACA	TAGACATCCT	2700
GCCCCATGGA	GACCAGACTC	AGATTGGGGA	ACGAGGCATC	AACTTGAGTA	CTGGTGGTCA	2760
GCGTCCAGAT	CAGTGTAGAC	CCGAGCCCTC	TACCAGCACA	CCAATGATTG	TCTTTTGGGA	2820
TGACCCTTTC	TCGGCTCTGG	ATGTCCATCT	GAGTGACCAC	CTAATGCAGG	CTGGCATCCT	2880
CGAGCTGCTC	CGGGATGACA	AGAGGACAGT	GGTCTTGGTG	ACCCACAAGC	TACAGTACCT	2940
GCCTCATGCT	GACTGGATCA	TTGCTATGAA	GGATGGCACC	ATTCAGAGGG	AGGGGACACT	3000
CAAGGACTTC	CAGAGGTCTG	AGTGCCAGCT	CTTTGAGCAT	TGGAAGACCC	TCATGAACCG	3060
GCAGGACCAA	GAGCTGGAGA	AGGAGACAGT	CATGGAGAGA	AAAGCCCCAG	AGCCATCTCA	3120
GGGCCTGCCC	CGTGCCATGT	CCTCAAGAGA	TGGCCTTCTG	CTGGATGAGG	ATGAGGAGGA	3180
AGAGGAGGCA	GCCGAGAGCG	AGGAAGATGA	CAACTTATCC	TCTGTGCTGC	ATCAGCGAGC	3240
CAAGATCCCA	TGGCGAGCCT	GCACCAAGTA	TTTGTCTCT	GCTGGCATCC	TGCTCCTGTC	3300
CCTGCTTGTC	TTCTCCAGC	TGCTCAAGCA	CATGGTCTTG	GTGGCCATTG	ACTACTGGCT	3360
GGCCAAGTGG	ACGGACAGTG	CCCTGGTCCT	GAGCCCCGCC	GCCAGGAAC	GCTCCCTCAG	3420
CCAGGAATGT	GCCCTGGACC	AATCTGTCTA	TGCCATGGTA	TTCACCGTGC	TCTGCAGCCT	3480
GGGTATCGCG	CTGTGCCTTG	TCACCTCTGT	CACTGTGGAG	TGGACGGGAC	TGAAGGTGGC	3540
CAAGAGGCTG	CATCGCAGCC	TGCTCAACCG	TATCATCCTG	GCTCCCATGA	GGTTCTTTGA	3600
GACCACGCCC	CTGGGGAGTA	TCCTGAACAG	ATTTTCATCT	GACTGTAAAC	CCATTGACCA	3660
GCATATCCCC	TCCACGCTGG	AGTGCCTGAG	CAGATCCACC	TACTCTGTG	TCTCCGCCCT	3720
GGCTGTCACT	TCCTACGTCA	CGCCTGTGTT	CCTAGTGGCC	CTCTTACCCC	TCGCCGTCGT	3780
GTGCTACTTC	ATCCAGAAGT	ACTTCCGAGT	GGCGTCCAGG	GACCTGCAGC	AGCTGGACGA	3840
CACAACACAG	CTCCCTCTGC	TCTCACACTT	TGCTGAAACT	GTGGAAGGAC	TCACCACCAT	3900
CCGTGCCTTC	AGGTACGAGG	CCCGGTTCCA	GCAGAAGCTC	CTAGAGTACA	CCGACTCCAA	3960
CAACATTGCC	TCTCTCTTCC	TCACAGCAGC	CAACAGGTGG	CTGGAAGTCC	GCATGGAGTA	4020
CATCGGAGCA	TGCGTGGTAC	TCATCGCCGC	TGCCACCTCC	ATCTCCAACT	CCCTACACAG	4080
GGAGCTCTCA	GCCGGCCTAG	TAGGCCTGGG	CCTCACCTAT	GCCTTGATGG	TCTCCAACTA	4140
CCTCAACTGG	ATGGTGAGGA	ACCTGGCAGA	CATGGAGATC	CAACTGGGAG	CTGTGAAGGG	4200
TATCCACACA	CTCCTGAAAA	CTGAGGCAGA	GAGCTATGAG	GGGCTCCTGG	CACCATCGCT	4260
GATCCCCAAG	AACTGGCCAG	ACCAAGGGAA	GATCCAAATT	CAAAACCTGA	GTGTACGCTA	4320
TGACAGCTCC	CTGAAGCCCG	TGCTGAAGCA	CGTCAACGCC	CTCATCTCCC	CAGGACAGAA	4380
GATTGGGATC	TGCGGCCGCA	CAGGCAGTGG	AAAATCCTCC	TTCTCTCTCG	CCTTTTTCCT	4440
AATGGTGGAT	ATGTTTGAAG	GGCGTATCAT	CATCGATGGC	ATTGACATCG	CCAAGCTGCC	4500

SUBSTITUTE SHEET (RULE 26)



GCTGCACACG	CTCGGCTCAC	GCCTGTCTAT	CATCCTACAG	GACCCTGTTT	TCTTCAGTGG	4560
TACCATCAGA	TTCAACCTGG	ACCCAGAGAA	GAAATGCTCA	GACAGCACGC	TGTGGGAGGC	4620
TCTGGAGATC	GCTCAGCTGA	AGCTGGTGGT	GAAGGCCCTG	CCAGGAGGCC	TGGATGCCAT	4680
CATCACGGAA	GGAGGGGAGA	ATTTTAGCCA	GGGCCAGAGG	CAGCTGTTCT	GCCTGGCCCC	4740
GGCCTTTGTG	AGGAAGACCA	GCATCTTCAT	CATGGATGAA	GCAACTGCCT	CCATCGACAT	4800
GGCTACGGAA	AATATCCTCC	AGAAGGTGGT	GATGACAGCC	TTCCGAGACC	GCACCGTGGT	4860
CACCATCGCG	CACCGCGTGC	ACACCATCCT	GAGTGCAGAC	CTAGTGATGG	TCCTGAAGAG	4920
GGGCGCGATC	CTGGAGTTCT	ACAAGCCGGA	AAAGCTTCTC	AGCCAGAAGG	ACAGCGTCTT	4980
TGCCTCCTTT	GTCCGCGCGG	ACAAATGACC	AGCCAGCGCC	AAAGTGCCAC	CCCACACCTC	5040
ACCTGCTTGC	CATGGATTTC	TTACTGTAAA	TCACTGTGAA	ATAAAGAAAC	TAATTCTTTG	5100
CTAAAAAAAA						5110

(32) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5110 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

- (A) NAME/KEY: CDS  
(B) LOCATION: 260..5004

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CCCTTGTGAC	AGGTCAGTCT	TACGAGAATA	TGGTAACTGA	GATCATGTCA	ATGGGCTATG		60
AACGAGAACA	AGTAATTGCA	GCCCTGAGAG	CCAGCTTCAA	CAACCCTGAT	AGAGCTGTGG		120
AATATCTTCT	AATGCGAATC	CCTGGAGACT	GAGGAGTTCC	AGTACTCACA	GCCTGTGGAG		180
GAGGATCAAC	CACGGCCTGA	CTTTCGCGGC	CGCCGCGGGA	GGCGCGCGGA	GCCGGAGCCG		240
AGCCCGTGCG	CGCGCCACC	ATG CCT TTG	GCC TTC TGC	GGC ACC GAG	AAC CAC		292
	Met Pro Leu	Ala Phe Cys	Gly Thr Glu	Asn His			
	1	5		10			
TCG GCC GCC	TAC CGG GTG	GAC CAA GGC	GTC CTC AAC	AAC GGC	TGC TTC		340
Ser Ala Ala	Tyr Arg Val	Asp Gln Gly	Val Leu Asn	Asn Gly	Cys Phe		
	15	20		25			
GTG GAC GCG	CTC AAT GTG	GTG CCA CAT	GTC TTT CTG	CTC TTC	ATC ACC		388
Val Asp Ala	Leu Asn Val	Val Pro His	Val Phe Leu	Leu Phe	Ile Thr		
	30	35		40			
TTC CCC ATC	CTC TTC ATC	GGA TGG GGC	AGC CAG AGC	TCC AAG GTG	CAC		436
Phe Pro Ile	Leu Phe Ile	Gly Trp Gly	Ser Ser Gln	Ser Ser Lys	Val His		
	45	50		55			

SUBSTITUTE SHEET (RULE 26)

ATT Ile 60	CAC His	CAC His	AGC Ser	ACC Thr	TGG Trp 65	CTC Leu	CAT His	TTC Phe	CCG Pro	GGG Gly 70	CAC His	AAC Asn	CTG Leu	CGC Arg	TGG Trp 75	484
ATC Ile	CTG Leu	ACC Thr	TTC Phe	ATA Ile 80	CTG Leu	CTC Leu	TTC Phe	GTC Val	CTC Leu	GTG Val	TGT Cys	GAG Glu	ATC Ile	GCT Ala	GAG Glu 90	532
GGT Gly	ATC Ile	CTG Leu	TCT Ser 95	GAC Asp	GGG Gly	GTG Val	ACA Thr	GAA Glu 100	TCC Ser	CGC Arg	CAC His	CTC Leu	CAC His	TTA Leu	TAC Tyr 105	580
ATG Met	CCA Pro	GCT Ala 110	GGG Gly	ATG Met	GCA Ala	TTC Phe	ATG Met	GCT Ala	GCC Ala	ATC Ile	ACC Thr	TCT Ser	GTG Val	GTC Val	TAC Tyr	628
TAC Tyr	CAT His	AAC Asn	ATT Ile	GAG Glu	ACC Thr	TCT Ser	AAC Asn	TTT Phe	CCC Pro	AAG Lys	CTG Leu	CTG Leu	ATT Ile	GCT Ala	CTG Leu	676
CTC Leu 140	ATC Ile	TAC Tyr	TGG Trp	ACC Thr	CTG Leu 145	GCC Ala	TTC Phe	ATC Ile	ACG Thr	AAG Lys 150	ACC Thr	ATC Ile	AAG Lys	TTC Phe	GTC Val 155	724
AAG Lys	TTC Phe	TAC Tyr	GAC Asp	CAC His 160	GCC Ala	ATT Ile	GGC Gly	TTC Phe	TCT Ser	CAG Gln 165	CTG Leu	CGC Arg	TTC Phe	TGC Cys	CTC Leu 170	772
ACG Thr	GGG Gly	CTT Leu	CTG Leu 175	GTG Val	ATC Ile	CTC Leu	TAC Tyr	GGG Gly 180	ATG Met	CTG Leu	CTG Leu	CTT Leu	GTG Val	GAG Glu	GTC Val 185	820
AAT Asn	GTC Val	ATC Ile	CGC Arg	GTG Val	AGG Arg	AGA Arg	TAC Tyr 195	GTC Val	TTC Phe	TTC Phe	AAG Lys	ACA Thr	CCA Pro	AGG Arg	GAA Glu	868
GTA Val	AAG Lys 205	CCC Pro	CCC Pro	GAG Glu	GAC Asp	CTA Leu 210	CAG Gln	GAC Asp	CTG Leu	GGT Gly	GTG Val	CGC Arg	TTT Phe	CTG Leu	CAG Gln	916
CCC Pro 220	TTC Phe	GTT Val	AAC Asn	CTG Leu	CTA Leu 225	TCA Ser	AAG Lys	GGG Gly	ACC Thr	TAC Tyr 230	TGG Trp	TGG Trp	ATG Met	AAT Asn	GCC Ala 235	964
TTC Phe	ATC Ile	AAG Lys	ACT Thr	GCT Ala 240	CAC His	AAG Lys	AAG Lys	CCC Pro	ATC Ile	GAC Asp 245	CTG Leu	CGG Arg	GCC Ala	ATC Ile	GGG Gly 250	1012
AAG Lys	CTG Leu	CCC Pro	ATT Ile 255	GCC Ala	ATG Met	AGA Arg	GCC Ala	CTC Leu 260	ACC Thr	AAC Asn	TAC Tyr	CAG Gln	CGA Arg	CTC Leu	TGC Cys 265	1060
TTG Leu	GCC Ala	TTC Phe 270	GAT Asp	GCC Ala	CAG Gln	GCG Ala	CGG Arg 275	AAG Lys	GAC Asp	ACA Thr	CAG Gln	AGC Ser	CAG Gln	CAG Gln	GGT Gly	1108
GCC Ala 285	CGG Ala	GCC Ala	ATC Ile	TGG Trp	AGG Arg	GCT Ala	CTC Leu 290	TGT Cys	CAT His	GCC Ala	TTT Phe 295	GGG Gly	AGA Arg	CGG Arg	CTG Leu	1156
GTC Val 300	CTC Leu	AGC Ser	AGC Ser	ACA Thr	TTC Phe 305	CGT Arg	ATC Ile	CTG Leu	GCC Ala	GAC Asp 310	CTC Leu	CTG Leu	GGC Gly	TTT Phe	GCT Ala 315	1204
GGG Gly	CCA Pro	CTC Leu	TGC Cys	ATC Ile 320	TTC Phe	GGG Gly	ATC Ile	GTG Val	GAC Asp 325	CAC His	CTC Leu	GGG Gly	AAG Lys	GAG Glu	AAC Asn 330	1252

SUBSTITUTE SHEET (RULE 26)

95

CAC GTC TTC CAG CCC AAG ACA CAG TTT CTT GGA GTT TAC TTT GTC TCA His Val Phe Gln Pro Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser 335 340 345	1300
TCC CAA GAG TTC CTC GGC AAT GCC TAT GTC TTG GCT GTT CTT CTG TTC Ser Gln Glu Phe Leu Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe 350 355 360	1348
CTT GCC CTC CTG CTG CAA AGG ACC TTT CTA CAA GCC TCG TAC TAC GTT Leu Ala Leu Leu Leu Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val 365 370 375	1396
GCC ATT GAA ACT GGG ATC AAC CTG AGA GGA GCA ATC CAG ACC AAG ATT Ala Ile Glu Thr Gly Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile 380 385 390 395	1444
TAC AAT AAG ATC ATG CAC TTG TCT ACT TCC AAC CTG TCC ATG GGG GAA Tyr Asn Lys Ile Met His Leu Ser Thr Ser Asn Leu Ser Met Gly Glu 400 405 410	1492
ATG ACT GCT GGG CAG ATC TGC AAC CTG GTG GCC ATC GAC ACC AAC CAG Met Thr Ala Gly Gln Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln 415 420 425	1540
CTC ATG TGG TTT TTC TTC TTA TGC CCA AAC CTC TGG GCT ATG CCG GTA Leu Met Trp Phe Phe Phe Leu Cys Pro Asn Leu Trp Ala Met Pro Val 430 435 440	1588
CAG ATC ATT GTG GGC GTG ATC CTC CTC TAC TAC ATC CTT GGG GTC AGC Gln Ile Ile Val Gly Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser 445 450 455	1636
GCC TTG ATT GGA GCG GCT GTC ATC ATT CTG CTG GCT CCT GTA CAG TAC Ala Leu Ile Gly Ala Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr 460 465 470 475	1684
TTT GTG GCC ACC AAG CTG TCC CAG GCA CAG CGG ACG ACC CTG GAA TAT Phe Val Ala Thr Lys Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu Tyr 480 485 490	1732
TCC AAT GAG AGG CTG AAG CAG ACC AAT GAG ATG CTC CGG GGC ATC AAG Ser Asn Glu Arg Leu Lys Gln Thr Asn Glu Met Leu Arg Gly Ile Lys 495 500 505	1780
TTG CTC AAG CTC TAT GCG TGG GAG AAC ATC TTC TGC TCC AGG GTG GAG Leu Leu Lys Leu Tyr Ala Trp Glu Asn Ile Phe Cys Ser Arg Val Glu 510 515 520	1828
AAG ACA CGC AGG AAG GAA ATG ACC AGC CTC AGG GCC TTC GCT GTC TAC Lys Thr Arg Arg Lys Glu Met Thr Ser Leu Arg Ala Phe Ala Val Tyr 525 530 535	1876
ACC TCC ATC TCC ATC TTC ATG AAC ACA GCT ATC CCC ATC GCT GCT GTC Thr Ser Ile Ser Ile Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val 540 545 550 555	1924
CTC ATC ACC TTC GTG GGC CAC GTC AGC TTC TTC AAA GAG TCG GAC TTC Leu Ile Thr Phe Val Gly His Val Ser Phe Phe Lys Glu Ser Asp Phe 560 565 570	1972
TCG CCC TCG GTG GCC TTT GCC TCT CTC TCT CTC TTC CAC ATC CTG GTC Ser Pro Ser Val Ala Phe Ala Ser Leu Ser Leu Phe His Ile Leu Val 575 580 585	2020
ACA CCG CTG TTC CTG CTG TCT AGT GTG GTT CGG TCC ACT GTC AAG GCC Thr Pro Leu Phe Leu Leu Ser Ser Val Val Arg Ser Thr Val Lys Ala 590 595 600	2068

SUBSTITUTE SHEET (RULE 26)



CTG GTG AGC GTG CAA AAG CTG AGT GAG TTC CTG TCC AGT GCA GAG ATC Leu Val Ser Val Gln Lys Leu Ser Glu Phe Leu Ser Ser Ala Glu Ile 605 610 615	2116
CGT GAG GAA CAG TGT GCC CCC CGA GAG CCC GCA CCC CAA GGC CAA GCG Arg Glu Glu Gln Cys Ala Pro Arg Glu Pro Ala Pro Gln Gly Gln Ala 620 625 630 635	2164
GGC AAG TAC CAG GCG GTG CCC CTC AAG GTC GTA AAC CGC AAG CGC CCA Gly Lys Tyr Gln Ala Val Pro Leu Lys Val Val Asn Arg Lys Arg Pro 640 645 650	2212
GCC CGA GAA GAA GTC CGG GAC CTC TTG GGC CCA CTG CAG AGG CTG ACT Ala Arg Glu Glu Val Arg Asp Leu Glu Gly Pro Leu Gln Arg Leu Thr 655 660 665	2260
CCC AGC ACG GAT GGA GAC GCT GAC AAC TTC TGT GTC CAG ATC ATC GGA Pro Ser Thr Asp Gly Asp Ala Asp Asn Phe Cys Val Gln Ile Ile Gly 670 675 680	2308
GGC TTC TTC ACC TGG ACC CCT GAT GGA ATC CCC ACC CTG TCC AAC ATC Gly Phe Phe Thr Trp Thr Pro Asp Gly Ile Pro Thr Leu Ser Asn Ile 685 690 695	2356
ACC ATC CGT ATC CCC CGA GGT CAG CTG ACC ATG ATC GTG GGG CAG GTG Thr Ile Arg Ile Pro Arg Gly Gln Leu Thr Met Ile Val Gly Gln Val 700 705 710 715	2404
GGC TGT GGC AAG TCC TCG CTC CTT CTG GCC ACC CTG GGG GAG ATG CAG Gly Cys Gly Lys Ser Ser Leu Leu Leu Ala Thr Leu Gly Glu Met Gln 720 725 730	2452
AAG GTC TCT GGA GCT GTC TTC TGG AAC AGC CTT CCA GAC AGC GAG GGG Lys Val Ser Gly Ala Val Phe Trp Asn Ser Leu Pro Asp Ser Glu Gly 735 740 745	2500
AGA AGA CCC CAG CAA CCC AGA GCG GGA GAC AGC GGC CGA TTC GGA TGC Arg Arg Pro Gln Gln Pro Arg Ala Gly Asp Ser Gly Arg Phe Gly Cys 750 755 760	2548
CAG GAG CAG AGG CCC TGT GGC TAC GCA TCT CAG AAA CCA TGG CTG CTA Gln Glu Gln Arg Pro Cys Gly Tyr Ala Ser Gln Lys Pro Trp Leu Leu 765 770 775	2596
AAT GCC ACT GTG GAG GAG AAC ATC ACC TTC GAG AGT CCC TTC AAT AAG Asn Ala Thr Val Glu Glu Asn Ile Thr Phe Glu Ser Pro Phe Asn Lys 780 785 790 795	2644
CAA CGG TAC AAG ATG GTC ATC GAA GCC TGC TCC CTG CAG CCA GAC ATA Gln Arg Tyr Lys Met Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile 800 805 810	2692
GAC ATC CTG CCC CAT GGA GAC CAG ACT CAG ATT GGG GAA CGA GGC ATC Asp Ile Leu Pro His Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile 815 820 825	2740
AAC TTG AGT ACT GGT GGT CAG CGT CCA GAT CAG TGT AGA CCC GAG CCC Asn Leu Ser Thr Gly Gly Gln Arg Pro Asp Gln Cys Arg Pro Glu Pro 830 835 840	2788
TCT ACC AGC ACA CCA ATG ATT GTC TTT TTG GAT GAC CCT TTC TCG GCT Ser Thr Ser Thr Pro Met Ile Val Phe Leu Asp Asp Pro Phe Ser Ala 845 850 855	2836
CTG GAT GTC CAT CTG AGT GAC CAC CTA ATG CAG GCT GGC ATC CTC GAG Leu Asp Val His Leu Ser Asp His Leu Met Gln Ala Gly Ile Leu Glu 860 865 870 875	2884

SUBSTITUTE SHEET (RULE 26)

CTG CTC CGG GAT GAC AAG AGG ACA GTG GTC TTG GTG ACC CAC AAG CTA	2932
Leu Leu Arg Asp Lys Arg Thr Val Val Leu Val Thr His Lys Leu	
880 885 890	
CAG TAC CTG CCT CAT GCT GAC TGG ATC ATT GCT ATG AAG GAT GGC ACC	2980
Gln Tyr Leu Pro His Ala Asp Trp Ile Ala Met Lys Asp Gly Thr	
895 900 905	
ATT CAG AGG GAG GGG ACA CTC AAG GAC TTC CAG AGG TCT GAG TGC CAG	3028
Ile Gln Arg Glu Gly Thr Leu Lys Asp Phe Gln Arg Ser Glu Cys Gln	
910 915 920	
CTC TTT GAG CAT TGG AAG ACC CTC ATG AAC CGG CAG GAC CAA GAG CTG	3076
Leu Phe Glu His Trp Lys Thr Leu Met Asn Arg Gln Asp Gln Glu Leu	
925 930 935	
GAG AAG GAG ACA GTC ATG GAG AGA AAA GCC CCA GAG CCA TCT CAG GGC	3124
Glu Lys Glu Thr Val Met Glu Arg Lys Ala Pro Glu Pro Ser Gln Gly	
940 945 950 955	
CTG CCC CGT GCC ATG TCC TCA AGA GAT GGC CTT CTG CTG GAT GAG GAT	3172
Leu Pro Arg Ala Met Ser Ser Arg Asp Gly Leu Leu Leu Asp Glu Asp	
960 965 970	
GAG GAG GAA GAG GAG GCA GCC GAG AGC GAG GAA GAT GAC AAC TTA TCC	3220
Glu Glu Glu Glu Glu Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser	
975 980 985	
TCT GTG CTG CAT CAG CGA GCC AAG ATC CCA TGG CGA GCC TGC ACC AAG	3268
Ser Val Leu His Gln Arg Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys	
990 995 1000	
TAT TTG TCC TCT GCT GGC ATC CTG CTC CTG TCC CTG CTT GTC TTC TCC	3316
Tyr Leu Ser Ser Ala Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser	
1005 1010 1015	
CAG CTG CTC AAG CAC ATG GTC TTG GTG GCC ATT GAC TAC TGG CTG GCC	3364
Gln Leu Leu Lys His Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala	
1020 1025 1030 1035	
AAG TGG ACG GAC AGT GCC CTG GTC CTG AGC CCC GCC GCC AGG AAC TGC	3412
Lys Trp Thr Asp Ser Ala Leu Val Leu Ser Pro Ala Ala Arg Asn Cys	
1040 1045 1050	
TCC CTC AGC CAG GAA TGT GCC CTG GAC CAA TCT GTC TAT GCC ATG GTA	3460
Ser Leu Ser Ser Gln Glu Cys Ala Leu Asp Gln Ser Val Tyr Ala Met Val	
1055 1060 1065	
TTC ACC GTG CTC TGC AGC CTG GGT ATC GCG CTG TGC CTT GTC ACC TCT	3508
Phe Thr Val Leu Cys Ser Leu Gly Ile Ala Leu Cys Leu Val Thr Ser	
1070 1075 1080	
GTC ACT GTG GAG TGG ACG GGA CTG AAG GTG GCC AAG AGG CTG CAT CGC	3556
Val Thr Val Glu Trp Thr Gly Leu Lys Val Ala Lys Arg Leu His Arg	
1085 1090 1095	
AGC CTG CTC AAC CGT ATC ATC CTG GCT CCC ATG AGG TTC TTT GAG ACC	3604
Ser Leu Leu Asn Arg Ile Ile Leu Ala Pro Met Arg Phe Phe Glu Thr	
1100 1105 1110 1115	
ACG CCC CTG GGG AGT ATC CTG AAC AGA TTT TCA TCT GAC TGT AAC ACC	3652
Thr Pro Leu Gly Ser Ile Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr	
1120 1125 1130	
ATT GAC CAG CAT ATC CCG TCC ACG CTG GAG TGC CTG AGC AGA TCC ACC	3700
Ile Asp Gln His Ile Pro Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr	
1135 1140 1145	

SUBSTITUTE SHEET (RULE 26)



TTA CTC TGT GTC TCC GCC CTG GCT GTC ATC TCC TAC GTC ACG CCT GTG Leu Leu Cys Val Ser Ala Leu Ala Val Ile Ser Tyr Val Thr Pro Val 1150 1155 1160	3748
TTC CTA GTG GCC CTC TTA CCC CTC GCC GTC GTG TGC TAC TTC ATC CAG Phe Leu Val Ala Leu Leu Pro Leu Ala Val Val Cys Tyr Phe Ile Gln 1165 1170 1175	3796
AAG TAC TTC CGA GTG GCG TCC AGG GAC CTG CAG CAG CTG GAC GAC ACA Lys Tyr Phe Arg Val Ala Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr 1180 1185 1190 1195	3844
ACA CAG CTC CCT CTG CTC TCA CAC TTT GCT GAA ACT GTG GAA GGA CTC Thr Gln Leu Pro Leu Leu Ser His Phe Ala Glu Thr Val Glu Gly Leu 1200 1205 1210	3892
ACC ACC ATC CGT GCC TTC AGG TAC GAG GCC CGG TTC CAG CAG AAG CTC Thr Thr Ile Arg Ala Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu 1215 1220 1225	3940
CTA GAG TAC ACC GAC TCC AAC AAC ATT GCC TCT CTC TTC CTC ACA GCA Leu Glu Tyr Thr Asp Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala 1230 1235 1240	3988
GCC AAC AGG TGG CTG GAA GTC CGC ATG GAG TAC ATC GGA GCA TGC GTG Ala Asn Arg Trp Leu Glu Val Arg Met Glu Tyr Ile Gly Ala Cys Val 1245 1250 1255	4036
GTA CTC ATC GCC GCT GCC ACC TCC ATC TCC AAC TCC CTA CAC AGG GAG Val Leu Ile Ala Ala Thr Ser Ile Ser Asn Ser Leu His Arg Glu 1260 1265 1270 1275	4084
CTC TCA GCC GGC CTA GTA GGC CTG GGC CTC ACC TAT GCC TTG ATG GTC Leu Ser Ala Gly Leu Val Gly Leu Gly Leu Thr Tyr Ala Leu Met Val 1280 1285 1290	4132
TCC AAC TAC CTC AAC TGG ATG GTG AGG AAC CTG GCA GAC ATG GAG ATC Ser Asn Tyr Leu Asn Trp Met Val Arg Asn Leu Ala Asp Met Glu Ile 1295 1300 1305	4180
CAA CTG GGA GCT GTG AAG GGT ATC CAC ACA CTC CTG AAA ACT GAG GCA Gln Leu Gly Ala Val Lys Gly Ile His Thr Leu Leu Lys Thr Glu Ala 1310 1315 1320	4228
GAG AGC TAT GAG GGG CTC CTG GCA CCA TCG CTG ATC CCC AAG AAC TGG Glu Ser Tyr Glu Gly Leu Leu Ala Pro Ser Leu Ile Pro Lys Asn Trp 1325 1330 1335	4276
CCA GAC CAA GGG AAG ATC CAA ATT CAA AAC CTG AGT GTA CGC TAT GAC Pro Asp Gln Gly Lys Ile Gln Ile Gln Asn Leu Ser Val Arg Tyr Asp 1340 1345 1350 1355	4324
AGC TCC CTG AAG CCC GTG CTG AAG CAC GTC AAC GCC CTC ATC TCC CCA Ser Ser Leu Lys Pro Val Leu Lys His Val Asn Ala Leu Ile Ser Pro 1360 1365 1370	4372
GGA CAG AAG ATT GGG ATC TGC GGC CGC ACA GGC AGT GGA AAA TCC TCC Gly Gln Lys Ile Gly Ile Cys Gly Arg Thr Gly Ser Gly Lys Ser Ser 1375 1380 1385	4420
TTC TCT CTC GCC TTT TTC CGA ATG GTG GAT ATG TTT GAA GGG CGT ATC Phe Ser Leu Ala Phe Phe Arg Met Val Asp Met Phe Glu Gly Arg Ile 1390 1395 1400	4468
ATC ATC GAT GGC ATT GAC ATC GCC AAG CTG CCG CTG CAC ACG CTC GGC Ile Ile Asp Gly Ile Asp Ile Ala Lys Leu Pro Leu His Thr Leu Gly 1405 1410 1415	4516

SUBSTITUTE SHEET (RULE 26)

99

TCA CGC CTG TCT ATC ATC CTA CAG GAC CCT GTT CTC TTC AGT GGT ACC 4564  
 Ser Arg Leu Ser Ile Ile Leu Gln Asp Pro Val Leu Phe Ser Gly Thr  
 1420 1425 1430 1435

ATC AGA TTC AAC CTG GAC CCA GAG AAG AAA TGC TCA GAC AGC ACG CTG 4612  
 Ile Arg Phe Asn Leu Asp Pro Glu Lys Lys Cys Ser Asp Ser Thr Leu  
 1440 1445 1450

TGG GAG GCT CTG GAG ATC GCT CAG CTG AAG CTG GTG GTG AAG GCC CTG 4660  
 Trp Glu Ala Leu Glu Ile Ala Gln Leu Lys Leu Val Val Lys Ala Leu  
 1455 1460 1465

CCA GGA GGC CTG GAT GCC ATC ATC ACG GAA GGA GGG GAG AAT TTT AGC 4708  
 Pro Gly Gly Leu Asp Ala Ile Ile Thr Glu Gly Gly Glu Asn Phe Ser  
 1470 1475 1480

CAG GGC CAG AGG CAG CTG TTC TGC CTG GCC CGG GCC TTT GTG AGG AAG 4756  
 Gln Gly Gln Arg Gln Leu Phe Cys Leu Ala Arg Ala Phe Val Arg Lys  
 1485 1490 1495

ACC AGC ATC TTC ATC ATG GAT GAA GCA ACT GCC TCC ATC GAC ATG GCT 4804  
 Thr Ser Ile Phe Ile Met Asp Glu Ala Thr Ala Ser Ile Asp Met Ala  
 1500 1505 1510 1515

ACG GAA AAT ATC CTC CAG AAG GTG GTG ATG ACA GCC TTC GCA GAC CGC 4852  
 Thr Glu Asn Ile Leu Gln Lys Val Val Met Thr Ala Phe Ala Asp Arg  
 1520 1525 1530

ACC GTG GTC ACC ATC GCG CAC CGC GTG CAC ACC ATC CTG AGT GCA GAC 4900  
 Thr Val Val Thr Ile Ala His Arg Val His Thr Ile Leu Ser Ala Asp  
 1535 1540 1545

CTA GTG ATG GTC CTG AAG AGG GGC GCG ATC CTG GAG TTC GAC AAG CCG 4948  
 Leu Val Met Val Leu Lys Arg Gly Ala Ile Leu Glu Phe Asp Lys Pro  
 1550 1555 1560

GAA AAG CTT CTC AGC CAG AAG GAC AGC GTC TTT GCC TCC TTT GTC CGC 4996  
 Glu Lys Leu Leu Ser Gln Lys Asp Ser Val Phe Ala Ser Phe Val Arg  
 1565 1570 1575

GCG GAC AA ATGACCAGCC AGCGCCAAAG TGCCACCCCA CACCTCACCT GCTTGCCATG 5054  
 Ala Asp  
 1580

GATTTCTTAC TGTAATCAC TTGTAATAA AGAACTAAT TCTTTGCTAA AAAAAA 5110

## (33) INFORMATION FOR SEQ ID NO:34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1581 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Pro Leu Ala Phe Cys Gly Thr Glu Asn His Ser Ala Ala Tyr Arg  
 1 5 10 15

Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn  
 20 25 30

Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe  
 35 40 45

SUBSTITUTE SHEET (RULE 26)

100

Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr  
 50 55 60  
 Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Ile  
 65 70 75 80  
 Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp  
 85 90 95  
 Gly Val Thr Glu Ser Arg His Leu His Leu Tyr Met Pro Ala Gly Met  
 100 105 110  
 Ala Phe Met Ala Ala Ile Thr Ser Val Val Tyr Tyr His Asn Ile Glu  
 115 120 125  
 Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr  
 130 135 140  
 Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Tyr Asp His  
 145 150 155 160  
 Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val  
 165 170 175  
 Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val  
 180 185 190  
 Arg Arg Tyr Val Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu  
 195 200 205  
 Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu  
 210 215 220  
 Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala  
 225 230 235 240  
 His Lys Lys Pro Ile Asp Leu Arg Ala Ile Gly Lys Leu Pro Ile Ala  
 245 250 255  
 Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Leu Ala Phe Asp Ala  
 260 265 270  
 Gln Ala Arg Lys Asp Thr Gln Ser Gln Gln Gly Ala Arg Ala Ile Trp  
 275 280 285  
 Arg Ala Leu Cys His Ala Phe Gly Arg Arg Leu Val Leu Ser Ser Thr  
 290 295 300  
 Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile  
 305 310 315 320  
 Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn His Val Phe Gln Pro  
 325 330 335  
 Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu  
 340 345 350  
 Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu  
 355 360 365  
 Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly  
 370 375 380  
 Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met  
 385 390 395 400  
 His Leu Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln

SUBSTITUTE SHEET (RULE 26)

101

405										410										415																																			
Ile	Cys	Asn	Leu	Val	Ala	Ile	Asp	Thr	Asn	Gln	Leu	Met	Trp	Phe	Phe																																								
			420					425					430																																										
Phe	Leu	Cys	Pro	Asn	Leu	Trp	Ala	Met	Pro	Val	Gln	Ile	Ile	Val	Gly																																								
		435					440					445																																											
Val	Ile	Leu	Leu	Tyr	Tyr	Ile	Leu	Gly	Val	Ser	Ala	Leu	Ile	Gly	Ala																																								
		450				455					460																																												
Ala	Val	Ile	Ile	Leu	Leu	Ala	Pro	Val	Gln	Tyr	Phe	Val	Ala	Thr	Lys																																								
		465			470					475					480																																								
Leu	Ser	Gln	Ala	Gln	Arg	Thr	Thr	Leu	Glu	Tyr	Ser	Asn	Glu	Arg	Leu																																								
				485					490					495																																									
Lys	Gln	Thr	Asn	Glu	Met	Leu	Arg	Gly	Ile	Lys	Leu	Leu	Lys	Leu	Tyr																																								
			500					505					510																																										
Ala	Trp	Glu	Asn	Ile	Phe	Cys	Ser	Arg	Val	Glu	Lys	Thr	Arg	Arg	Lys																																								
		515					520					525																																											
Glu	Met	Thr	Ser	Leu	Arg	Ala	Phe	Ala	Val	Tyr	Thr	Ser	Ile	Ser	Ile																																								
		530				535					540																																												
Phe	Met	Asn	Thr	Ala	Ile	Pro	Ile	Ala	Ala	Val	Leu	Ile	Thr	Phe	Val																																								
		545			550				555					560																																									
Gly	His	Val	Ser	Phe	Phe	Lys	Glu	Ser	Asp	Phe	Ser	Pro	Ser	Val	Ala																																								
				565					570					575																																									
Phe	Ala	Ser	Leu	Ser	Leu	Phe	His	Ile	Leu	Val	Thr	Pro	Leu	Phe	Leu																																								
			580					585				590																																											
Leu	Ser	Ser	Val	Val	Arg	Ser	Thr	Val	Lys	Ala	Leu	Val	Ser	Val	Gln																																								
		595				600					605																																												
Lys	Leu	Ser	Glu	Phe	Leu	Ser	Ser	Ala	Glu	Ile	Arg	Glu	Glu	Gln	Cys																																								
		610				615					620																																												
Ala	Pro	Arg	Glu	Pro	Ala	Pro	Gln	Gly	Gln	Ala	Gly	Lys	Tyr	Gln	Ala																																								
		625			630				635					640																																									
Val	Pro	Leu	Lys	Val	Val	Asn	Arg	Lys	Arg	Pro	Ala	Arg	Glu	Glu	Val																																								
				645					650					655																																									
Arg	Asp	Leu	Leu	Gly	Pro	Leu	Gln	Arg	Leu	Thr	Pro	Ser	Thr	Asp	Gly																																								
			660					665					670																																										
Asp	Ala	Asp	Asn	Phe	Cys	Val	Gln	Ile	Ile	Gly	Gly	Phe	Phe	Thr	Trp																																								
		675					680					685																																											
Thr	Pro	Asp	Gly	Ile	Pro	Thr	Leu	Ser	Asn	Ile	Thr	Ile	Arg	Ile	Pro																																								
		690				695					700																																												
Arg	Gly	Gln	Leu	Thr	Met	Ile	Val	Gly	Gln	Val	Gly	Cys	Gly	Lys	Ser																																								
		705			710				715					720																																									
Ser	Leu	Leu	Leu	Ala	Thr	Leu	Gly	Glu	Met	Gln	Lys	Val	Ser	Gly	Ala																																								
				725				730					735																																										
Val	Phe	Trp	Asn	Ser	Leu	Pro	Asp	Ser	Glu	Gly	Arg	Arg	Pro	Gln	Gln																																								
			740					745					750																																										
Pro	Arg	Ala	Gly	Asp	Ser	Gly	Arg	Phe	Gly	Cys	Gln	Glu	Gln	Arg	Pro																																								
		755					760					765																																											

SUBSTITUTE SHEET (RULE 26)

102

Cys Gly Tyr Ala Ser Gln Lys Pro Trp Leu Leu Asn Ala Thr Val Glu  
 770 775 780  
 Glu Asn Ile Thr Phe Glu Ser Pro Phe Asn Lys Gln Arg Tyr Lys Met  
 785 790 795 800  
 Val Ile Glu Ala Cys Ser Leu Gln Pro Asp Ile Asp Ile Leu Pro His  
 805 810 815  
 Gly Asp Gln Thr Gln Ile Gly Glu Arg Gly Ile Asn Leu Ser Thr Gly  
 820 825 830  
 Gly Gln Arg Pro Asp Gln Cys Arg Pro Glu Pro Ser Thr Ser Thr Pro  
 835 840 845  
 Met Ile Val Phe Leu Asp Asp Pro Phe Ser Ala Leu Asp Val His Leu  
 850 855 860  
 Ser Asp His Leu Met Gln Ala Gly Ile Leu Glu Leu Leu Arg Asp Asp  
 865 870 875 880  
 Lys Arg Thr Val Val Leu Val Thr His Lys Leu Gln Tyr Leu Pro His  
 885 890 895  
 Ala Asp Trp Ile Ile Ala Met Lys Asp Gly Thr Ile Gln Arg Glu Gly  
 900 905 910  
 Thr Leu Lys Asp Phe Gln Arg Ser Glu Cys Gln Leu Phe Glu His Trp  
 915 920 925  
 Lys Thr Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val  
 930 935 940  
 Met Glu Arg Lys Ala Pro Glu Pro Ser Gln Gly Leu Pro Arg Ala Met  
 945 950 955 960  
 Ser Ser Arg Asp Gly Leu Leu Leu Asp Glu Asp Glu Glu Glu Glu Glu  
 965 970 975  
 Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Val Leu His Gln  
 980 985 990  
 Arg Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala  
 995 1000 1005  
 Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His  
 1010 1015 1020  
 Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser  
 1025 1030 1035 1040  
 Ala Leu Val Leu Ser Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu  
 1045 1050 1055  
 Cys Ala Leu Asp Gln Ser Val Tyr Ala Met Val Phe Thr Val Leu Cys  
 1060 1065 1070  
 Ser Leu Gly Ile Ala Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp  
 1075 1080 1085  
 Thr Gly Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg  
 1090 1095 1100  
 Ile Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser  
 1105 1110 1115 1120  
 Ile Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile

SUBSTITUTE SHEET (RULE 26)

103

1125	1130	1135
Pro Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser 1140	1145	1150
Ala Leu Ala Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu 1155	1160	1165
Leu Pro Leu Ala Val Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val 1170	1175	1180
Ala Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu 1185	1190	1195
Leu Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala 1205	1210	1215
Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp 1220	1225	1230
Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu 1235	1240	1245
Glu Val Arg Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala 1250	1255	1260
Ala Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu 1265	1270	1275
Val Gly Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn 1285	1290	1295
Trp Met Val Arg Asn Leu Ala Asp Met Glu Ile Gln Leu Gly Ala Val 1300	1305	1310
Lys Gly Ile His Thr Leu Leu Lys Thr Glu Ala Glu Ser Tyr Glu Gly 1315	1320	1325
Leu Leu Ala Pro Ser Leu Ile Pro Lys Asn Trp Pro Asp Gln Gly Lys 1330	1335	1340
Ile Gln Ile Gln Asn Leu Ser Val Arg Tyr Asp Ser Ser Leu Lys Pro 1345	1350	1355
Val Leu Lys His Val Asn Ala Leu Ile Ser Pro Gly Gln Lys Ile Gly 1365	1370	1375
Ile Cys Gly Arg Thr Gly Ser Gly Lys Ser Ser Phe Ser Leu Ala Phe 1380	1385	1390
Phe Arg Met Val Asp Met Phe Glu Gly Arg Ile Ile Ile Asp Gly Ile 1395	1400	1405
Asp Ile Ala Lys Leu Pro Leu His Thr Leu Gly Ser Arg Leu Ser Ile 1410	1415	1420
Ile Leu Gln Asp Pro Val Leu Phe Ser Gly Thr Ile Arg Phe Asn Leu 1425	1430	1435
Asp Pro Glu Lys Lys Cys Ser Asp Ser Thr Leu Trp Glu Ala Leu Glu 1445	1450	1455
Ile Ala Gln Leu Lys Leu Val Val Lys Ala Leu Pro Gly Gly Leu Asp 1460	1465	1470
Ala Ile Ile Thr Glu Gly Gly Glu Asn Phe Ser Gln Gly Gln Arg Gln 1475	1480	1485

SUBSTITUTE SHEET (RULE 26)

104

Leu Phe Cys Leu Ala Arg Ala Phe Val Arg Lys Thr Ser Ile Phe Ile  
 1490 1495 1500  
 Met Asp Glu Ala Thr Ala Ser Ile Asp Met Ala Thr Glu Asn Ile Leu  
 1505 1510 1515 1520  
 Gln Lys Val Val Met Thr Ala Phe Ala Asp Arg Thr Val Val Thr Ile  
 1525 1530 1535  
 Ala His Arg Val His Thr Ile Leu Ser Ala Asp Leu Val Met Val Leu  
 1540 1545 1550  
 Lys Arg Gly Ala Ile Leu Glu Phe Asp Lys Pro Glu Lys Leu Leu Ser  
 1555 1560 1565  
 Gln Lys Asp Ser Val Phe Ala Ser Phe Val Arg Ala Asp  
 1570 1575 1580

## (34) INFORMATION FOR SEQ ID NO:35:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 4877 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AGCCGAGCCC GTGCGCGCGC CGCCATGCCC TTGGCCTTCT GCGGTACCGA GAACCACTCG	60
GCCGCCTACC GGGTGGACCA GGGCGTCTC AACAAACGGCT GCTTCGTGGA CGCGCTCAAC	120
GTGGTGCCGC ACGTTTTCTT GCTCTTCATC ACCTTCCCCA TCCTCTTCAT CGGATGGGGC	180
AGCCAGAGCT CCAAGGTGCA CATCCACCAC AGCACCTGGC TGCACCTTCC AGGGCACAAC	240
CTGCGCTGGA TCCTTACCTT CATTTTGCTC TTCGTCTTG TGTGTGAGAT CGCTGAGGGC	300
ATCCTGTCTG ATGGGGTGAC AGAATCCCGC CACCTCCACC TGTACATGCC AGCCGGGATG	360
GCGTTCATGG CTGCCATCAC CTCTGTAGTC TACTATCATA ACATCGAGAC CTCCAACTTC	420
CCCAAGCTTT TGATCGCTCT GTCATCTAT TGGACCCTGG CCTTCATCAC GAAGACCATC	480
AAGTTTGTC AATTCTATGA CCACGCCATC GGCTTCTCCC AGCTGCGCTT CTGCCTCACG	540
GGGCTTCTGG TGATCCTGTA TGGGATGTTG CTGCTTGTGG AGGTCAACGT CATCAGAGTG	600
AGGAGGTACA TCTTCTTCAA GACGCCACGG GAGGTGAAGC CCCCTGAGGA CCTGCAGGAC	660
CTGGGTGTGC GCTTTCTGCA GCCCTTCGTT AACCTGCTGT CAAAGGGGAC CTATTGGTGG	720
ATGAATGCCT TCATCAAGAC GGCCCAAG AAGCCCATCG ACCTGCGGGC CATCGCGAAG	780
CTGCCCATCG CCATGAGAGC CCTCACCAAC TATCAGCGCC TCTGCGTGGC CTTCGATGCT	840
CAGGCGCGGA AGGACACACA GAGCCACAG GGTGCCCGGG CCATCTGGAG GGCTCTATGC	900
CATGCCTTTG GGAGACGCCT GATCCTCAGC AGCACATTCC GCATCCTGGC TGACCTGTTG	960

SUBSTITUTE SHEET (RULE 26)

GGCTTCGCTG	GACCACTCTG	CATCTTTGGG	ATCGTGGACC	ACCTGGGGAA	GGAGAACCAC	1020
GTCTTCCAGC	CCAAGACACA	GTTTCTCGGG	GTTTACTTCG	TCTCTTCTCA	AGAGTTCCCTT	1080
GGCAATGCCT	ACGTCTTGGC	CGTGCTTCTG	TTCCTTGCCC	TGCTACTGCA	AAGGACATTCT	1140
CTGCAAGCCT	CATACTACGT	CGCCATTGAA	ACTGGAATTA	ACCTGAGAGG	AGCAATCCAG	1200
ACCAAGATTT	ACAATAAAAT	CATGCACATG	TCCACCTCCA	ACCTGTCAAT	GGGGGAAATG	1260
ACTGTGGGC	AGATCTGCAA	CCTGGTGGCC	ATCGACACAA	ACCAGCTCAT	GTGGTTCTTC	1320
TTTCTGTGCC	CAAACCTCTG	GACGATGCCA	GTACAGATCA	TTGTGGGCGT	GATCCTTCTC	1380
TACTACATCC	TTGGGGTCAG	TGCCTTGATT	GGAGCAGCTG	TCATCATTCT	GCTGGCTCCT	1440
GTACAGTACT	TTGTGGCCAC	CAAGCTCTCC	CAGGCACAGC	GGACGACCTT	GGAGCACTCC	1500
AACGAGAGGC	TGAAGCAGAC	CAACGAGATG	CTCCGGGGCA	TGAAGCTGCT	CAAACGTAT	1560
GCGTGGGAGA	GCATCTTCTG	CTCCAGGGTG	GAGGTGACTC	GCAGGAAGGA	GATGACCAGC	1620
CTGAGGGCGT	TTGCTGTCTA	CACTTCCATC	TCCATCTTCA	TGAACACAGC	CATCCCCATT	1680
GCTGCCGTCC	TCATCACCTT	CGTGGGCCAC	GTCAGCTTCT	TCAAAGAGTC	GGACTTGTCA	1740
CCCTCGGTGG	CCTTGCCTC	CCTCTCTCTC	ITCCACATCC	TGGTCACTCC	ACTGTTCCCTG	1800
CTGTCTAGCG	TGGTTCGGTC	CACTGTCAAA	GCCCTGGTGA	GCGTGCAAAA	ACTGAGCGAG	1860
TTCTGTCTA	GTGCAGAGAT	CCGTGAGGAG	CAGTGTGCCC	CCCGAGAGCC	TGCACCCCAA	1920
GGCCAAGCCG	GCAAGTACCA	GGCAGTGCCC	CTCAAGGTTG	TGAACCGCAA	ACGCCAGCC	1980
CGGGAAGAGG	TCCGGGACCT	CCTGGGCCCA	CTGCAGAGGC	TGGCCCCTAG	CATGGACGGG	2040
GATGCTGACA	ACTTCTGTGT	CCAGATCATC	GGAGGCTTCT	TCACCTGGAC	CCCTGATGGA	2100
ATCCCCACTC	TGTCCAACAT	CACCATCCGT	ATTCCCCGAG	GTCAGCTAAC	CATGATTGTG	2160
GGGCAGGTGG	GCTGCGGCAA	GTCCTCGCTC	CTCCTCGCCA	CCCTGGGGGA	GATGCAGAAG	2220
GTGTGCGGGG	CCGTCTTCTG	GAACAGCAAC	CTTCCGGACA	GCGAGGGGAG	AGGACCCAG	2280
CAGCCCAGAG	CGGGAGACAG	CAGCTGGCTC	GGATATCAGG	AGCAGAGGCC	CCGTGGCTAC	2340
GCATCTCAGA	AACCATGGCT	GCTAAACGCC	ACCGTGGAAG	AGAACATCAC	CTTCGAGAGT	2400
CCCTTCAATC	CGCAGCGGTA	CAAGATGGTC	ATCGAAGCCT	GCTCCCTGCA	GCCGGACATA	2460
GACATCCTGC	CCCACGGAGA	CCAGACTCAG	ATTGGGGAAC	GGGGCATCAA	CCTGTCTGGT	2520
GGTCAGCGTC	CAGATCAGTG	TGGTCCAGAG	CCCTCTACCA	GCAGACCAAT	GTTTGTCTTC	2580
TTGGATGACC	CCTTCTCAGC	TTTGGATGTC	CATCTGAGTG	ACCACCTGAT	GCAGGCCGGC	2640
ATCCTTGAGC	TGCTCCGGGA	TGACAAGAGG	ACAGTGGTCT	TGGTGACCCA	CAAGCTACAG	2700
TATCTGCCTC	ATGCAGACTG	GATCATTGCC	ATGAAGGATG	GGACCATTC	GAGGGAAGGG	2760
ACGCTCAAGG	ACTTCCAGAG	GTCCGAGTGC	CAGCTCTTTG	AGCACTGGAA	GACCTCATG	2820
AACCGGCAGG	ACCAAGAGCT	GGAGAAGGAG	ACAGTCATGG	AGAGGAAAGC	CTCAGAGCCA	2880
TCTCAGGGCC	TGCCCCGTGC	CATGTCTCTC	AGAGACGGCC	TTCTGCTGGA	TGAGGAAGAG	2940
GAGGAAGAGG	AGGCAGCCGA	AAGCGAGGAA	GATGACAACT	TATCTTCAGT	GCTGCATCAG	3000

SUBSTITUTE SHEET (RULE 26)



CGAGCTAAGA	TCCCCTGGCG	AGCCTGCACT	AAGTATCTGT	CCTCTGCTGG	CATTCTGCTC	3060
CTGTCCCTGC	TTGTCTTCTC	CCAGCTGCTC	AAGCACATGG	TCTTGGTGGC	CATTGATTAT	3120
TGGCTGGCCA	AGTGGACGGA	CAGTGGCCTG	GTCCTGAGCC	CCGCTGCCAG	GAAGTGTTCG	3180
CTCAGCCAGG	AATGTGACCT	GGACCACTCT	GTCTATGCCA	TGGTATTAC	CTTGCTCTGC	3240
AGCCTGGGTA	TCGTGCTGTG	CCTGGTCACC	TCTGTCACTG	TGGAGTGGAC	GGGACTGAAG	3300
GTGGCCAAGA	GGCTACACCG	CAGCCTGCTC	AACCGCATCA	TCCTGGCCCC	CATGAGGTTT	3360
TTTGAGACCA	CACCCCTCGG	GAGTATCCTG	AACAGATTTT	CATCCGACTG	TAACACCATT	3420
GACCAGCACA	TCCCATCCAC	GCTGGAGTGT	CTGAGCCGGT	CCACCCTGCT	GTGTGTCTCC	3480
GCCCTGACTG	TCATCTCCTA	TGTCACACCC	GTGTTCCCTG	TGGCCCTCTT	ACCCCTAGCT	3540
GTTGTGTGCT	ACTTCATTCA	GAAGTACTTC	CGAGTGGCAT	CCAGGGACCT	GCAGCAGCTG	3600
GACGACACGA	CGCAGCTCCC	GCTCGTCTCA	CACTTTGCTG	AAACTGTGGA	GGGACTCACC	3660
ACCATCCGTG	CCTTCAGGTA	CGAGGCCCGG	TTCCAGCAGA	AGCTTCTAGA	ATATACCGAC	3720
TCCAACAACA	TCGCCTCCCT	CTTCCTCAGC	GCAGCCAACA	GATGGCTGGA	AGTCTGCATG	3780
GAGTACATCG	GAGCGTGCGT	GGTACTCATT	GCGGCTGCCA	CCTCCATCTC	CAACTCCCTG	3840
CACAGGGAAC	TTTCTGTGCG	CCTGGTGGGC	CTGGGCCTCA	CCTATGCCTT	GATGGTCTCC	3900
AACTACCTCA	ACTGGATGGT	GAGGAACCTG	GCGGACATGG	AGATCCAGCT	GGGGGCTGTG	3960
AAGAGGATCC	ACGCACTCCT	GAAAACCGAG	GCGGAGAGCT	ATGAGGGGCT	CCTGGCGCCG	4020
TCGTTGATCC	CCAAGAACTG	GCCAGACCAA	GGGAAGATCC	AAATTGAGAA	CCTGAGCGTG	4080
CGCTATGACA	GCTCCCTGAA	GCCAGTGCTG	AAGCATGTCA	ACACCCTCAT	CTCCCCGGGG	4140
CAGAAGATCG	GGATCTGCGG	CCGCACAGGC	AGCGGGAAGT	CCTCCTTCTC	CCTGGCCCTT	4200
TTCCGAATGG	TGGACATGTT	TGAAGGACGC	ATCATCATTG	ATGGCATCGA	CATCGCCAAG	4260
CTGCCACTTC	ACACGCTGCG	CTCACGCCTG	TCCATCATCC	TACAGGACCC	CGTCCTCTTC	4320
AGCGGCACGA	TCAGATTCAA	CCTGGACCCC	GAGAAGAAAT	GCTCAGACAG	CACACTGTGG	4380
GAGGCCCTGG	AGATCGCCCC	GCTGAAGCTG	GTAGTGAAGG	CACTGCCAGG	AGGCCTAGAT	4440
GCCATCATCA	CAGAAGGAGG	GGAGAATTTT	AGCCAGGGCC	AGAGGCAGCT	GTTCTGCCTG	4500
GCCCCGGCCT	TCGTGAGGAA	GACCAGCATC	TTCATCATGG	ATGAAGCAAC	CGCCTCCATC	4560
GACATGGCTA	CGGAGAACAT	CCTCCAGAAG	GTGGTGATGA	CAGCCTTCGC	AGACCGCACG	4620
GTGGTCACCA	TCGCGCATCG	TGTGCACACC	ATCCTGAGTG	CAGACCTGGT	GATGGTCCTC	4680
AAGAGGGGTG	CTATCCTGGA	GTTTGACAAG	CCAGAGACGC	TCCTCAGCCA	GAAGGACAGC	4740
GTGTTGCGCT	CCTTTGTCCG	TGCGGACAAG	TGACTTACCG	GAGCCAAAGT	GCCACCCCGC	4800
GCCTCGCTTG	CTTGCCCTAGG	ATTTCTAACT	GCAAATCACT	TGTAAATAAA	TTAATTCTTT	4860
GCTAAAAAAA	AAAAAAA					4877

(35) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

SUBSTITUTE SHEET (RULE 26)

107

(A) LENGTH: 4877 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 25..4770

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

AGCCGAGCCC GTGCGCGCGC CGCC ATG CCC TTG GCC TTC TGC GGT ACC GAG	51
Met Pro Leu Ala Phe Cys Gly Thr Glu	
1 5	
AAC CAC TCG GCC GCC TAC CGG GTG GAC CAG GGC GTC CTC AAC AAC GGC	99
Asn His Ser Ala Ala Tyr Arg Val Asp Gln Gly Val Leu Asn Asn Gly	
10 15 20 25	
TGC TTC GTG GAC GCG CTC AAC GTG GTG CCG CAC GTT TTC CTG CTC TTC	147
Cys Phe Val Asp Ala Leu Asn Val Val Pro His Val Phe Leu Leu Phe	
30 35 40	
ATC ACC TTC CCC ATC CTC TTC ATC GGA TGG GGC AGC CAG AGC TCC AAG	195
Ile Thr Phe Pro Ile Leu Phe Ile Gly Trp Gly Ser Gln Ser Ser Lys	
45 50 55	
GTG CAC ATC CAC CAC AGC ACC TGG CTG CAC TTT CCA GGG CAC AAC CTG	243
Val His Ile His His Ser Thr Trp Leu His Phe Pro Gly His Asn Leu	
60 65 70	
CGC TGG ATC CTT ACC TTC ATT TTG CTC TTC GTC CTT GTG TGT GAG ATC	291
Arg Trp Ile Leu Thr Phe Ile Leu Leu Phe Val Leu Val Cys Glu Ile	
75 80 85	
GCT GAG GGC ATC CTG TCT GAT GGG GTG ACA GAA TCC CGC CAC CTC CAC	339
Ala Glu Gly Ile Leu Ser Asp Gly Val Thr Glu Ser Arg His Leu His	
90 95 100 105	
CTG TAC ATG CCA GCC GGG ATG GCG TTC ATG GCT GCC ATC ACC TCT GTA	387
Leu Tyr Met Pro Ala Gly Met Ala Phe Met Ala Ala Ile Thr Ser Val	
110 115 120	
GTC TAC TAT CAT AAC ATC GAG ACC TCC AAC TTC CCC AAG CTT TTG ATC	435
Val Tyr Tyr His Asn Ile Glu Thr Ser Asn Phe Pro Lys Leu Leu Ile	
125 130 135	
GCT CTG CTC ATC TAT TGG ACC CTG GCC TTC ATC ACG AAG ACC ATC AAG	483
Ala Leu Leu Ile Tyr Trp Thr Leu Ala Phe Ile Thr Lys Thr Ile Lys	
140 145 150	
TTT GTC AAG TTC TAT GAC CAC GCC ATC GGC TTC TCC CAG CTG CGC TTC	531
Phe Val Lys Phe Tyr Asp His Ala Ile Gly Phe Ser Gln Leu Arg Phe	
155 160 165	
TGC CTC ACG GGG CTT CTG GTG ATC CTG TAT GGG ATG TTG CTG CTT GTG	579
Cys Leu Thr Gly Leu Leu Val Ile Leu Tyr Gly Met Leu Leu Leu Val	
170 175 180 185	

SUBSTITUTE SHEET (RULE 26)

108

GAG	GTC	AAC	GTC	ATC	AGA	GTG	AGG	AGG	TAC	ATC	TTC	TTC	AAG	ACG	CCA	627
Glu	Val	Asn	Val	Ile	Arg	Val	Arg	Arg	Tyr	Ile	Phe	Phe	Lys	Thr	Pro	
				190					195					200		
CGG	GAG	GTG	AAG	CCC	CCT	GAG	GAC	CTG	CAG	GAC	CTG	GGT	GTG	CGC	TTT	675
Arg	Glu	Val	Lys	Pro	Pro	Glu	Asp	Leu	Gln	Asp	Leu	Gly	Val	Arg	Phe	
			205					210					215			
CTG	CAG	CCC	TTC	GTT	AAC	CTG	CTG	TCA	AAG	GGG	ACC	TAT	TGG	TGG	ATG	723
Leu	Gln	Pro	Phe	Val	Asn	Leu	Leu	Ser	Lys	Gly	Thr	Tyr	Trp	Trp	Met	
			220					225					230			
AAT	GCC	TTC	ATC	AAG	ACG	GCC	CAC	AAG	AAG	CCC	ATC	GAC	CTG	CGG	GCC	771
Asn	Ala	Phe	Ile	Lys	Thr	Ala	His	Lys	Lys	Pro	Ile	Asp	Leu	Arg	Ala	
			235				240				245					
ATC	GCG	AAG	CTG	CCC	ATC	GCC	ATG	AGA	GCC	CTC	ACC	AAC	TAT	CAG	CGC	819
Ile	Ala	Lys	Leu	Pro	Ile	Ala	Met	Arg	Ala	Leu	Thr	Asn	Tyr	Gln	Arg	
					255					260					265	
CTC	TGC	GTG	GCC	TTC	GAT	GCT	CAG	GCG	CGG	AAG	GAC	ACA	CAG	AGC	CCA	867
Leu	Cys	Val	Ala	Phe	Asp	Ala	Gln	Ala	Arg	Lys	Asp	Thr	Gln	Ser	Pro	
				270					275						280	
CAG	GGT	GCC	CGG	GCC	ATC	TGG	AGG	GCT	CTA	TGC	CAT	GCC	TTT	GGG	AGA	915
Gln	Gly	Ala	Arg	Ala	Ile	Trp	Arg	Ala	Leu	Cys	His	Ala	Phe	Gly	Arg	
			285					290					295			
CGC	CTG	ATC	CTC	AGC	AGC	ACA	TTC	CGC	ATC	CTG	GCT	GAC	CTG	TTG	GGC	963
Arg	Leu	Ile	Leu	Ser	Ser	Thr	Phe	Arg	Ile	Leu	Ala	Asp	Leu	Leu	Gly	
			300				305						310			
TTC	GCT	GGA	CCA	CTC	TGC	ATC	TTT	GGG	ATC	GTG	GAC	CAC	CTG	GGG	AAG	1011
Phe	Ala	Gly	Pro	Leu	Cys	Ile	Phe	Gly	Ile	Val	Asp	His	Leu	Gly	Lys	
			315				320					325				
GAG	AAC	CAC	GTC	TTC	CAG	CCC	AAG	ACA	CAG	TTT	CTC	GGG	GTT	TAC	TTC	1059
Glu	Asn	His	Val	Phe	Gln	Pro	Lys	Thr	Gln	Phe	Leu	Gly	Val	Tyr	Phe	
					335					340					345	
GTC	TCT	TCT	CAA	GAG	TTC	CTT	GGC	AAT	GCC	TAC	GTC	TTG	GCC	GTG	CTT	1107
Val	Ser	Ser	Gln	Glu	Phe	Leu	Gly	Asn	Ala	Tyr	Val	Leu	Ala	Val	Leu	
				350					355						360	
CTG	TTC	CTT	GCC	CTG	CTA	CTG	CAA	AGG	ACA	TTC	CTG	CAA	GCC	TCA	TAC	1155
Leu	Phe	Leu	Ala	Leu	Leu	Leu	Gln	Arg	Thr	Phe	Leu	Gln	Ala	Ser	Tyr	
			365					370					375			
TAC	GTC	GCC	ATT	GAA	ACT	GGA	ATT	AAC	CTG	AGA	GGA	GCA	ATC	CAG	ACC	1203
Tyr	Val	Ala	Ile	Glu	Thr	Gly	Ile	Asn	Leu	Arg	Gly	Ala	Ile	Gln	Thr	
			380				385						390			
AAG	ATT	TAC	AAT	AAA	ATC	ATG	CAC	ATG	TCC	ACC	TCC	AAC	CTG	TCA	ATG	1251
Lys	Ile	Tyr	Asn	Lys	Ile	Met	His	Met	Ser	Thr	Ser	Asn	Leu	Ser	Met	
			395				400					405				
GGG	GAA	ATG	ACT	GCT	GGG	CAG	ATC	TGC	AAC	CTG	GTG	GCC	ATC	GAC	ACA	1299
Gly	Glu	Met	Thr	Ala	Gly	Gln	Ile	Cys	Asn	Leu	Val	Ala	Ile	Asp	Thr	
					415					420					425	
AAC	CAG	CTC	ATG	TGG	TTC	TTC	TTT	CTG	TGC	CCA	AAC	CTC	TGG	ACG	ATG	1347
Asn	Gln	Leu	Met	Trp	Phe	Phe	Phe	Leu	Cys	Pro	Asn	Leu	Trp	Thr	Met	
				430					435						440	
CCA	GTA	CAG	ATC	ATT	GTG	GGC	GTG	ATC	CTT	CTC	TAC	TAC	ATC	CTT	GGG	1395
Pro	Val	Gln	Ile	Ile	Val	Gly	Val	Ile	Leu	Leu	Tyr	Tyr	Ile	Leu	Gly	
			445					450							455	

SUBSTITUTE SHEET (RULE 26)

109

GTC	AGT	GCC	TTG	ATT	GGA	GCA	GCT	STC	ATC	ATT	CTG	CTG	GCT	CCT	GTA	1443
Val	Ser	Ala	Leu	Ile	Gly	Ala	Ala	Val	Ile	Ile	Leu	Leu	Ala	Pro	Val	
		460					465					470				
CAG	TAC	TTT	GTG	GCC	ACC	AAG	CTC	TCC	CAG	GCA	CAG	CGG	ACG	ACC	TTG	1491
Gln	Tyr	Phe	Val	Ala	Thr	Lys	Leu	Ser	Gln	Ala	Gln	Arg	Thr	Thr	Leu	
		475				480					485					
GAG	CAC	TCC	AAC	GAG	AGG	CTG	AAG	CAG	ACC	AAC	GAG	ATG	CTC	CGG	GGC	1539
Glu	His	Ser	Asn	Glu	Arg	Leu	Lys	Gln	Thr	Asn	Glu	Met	Leu	Arg	Gly	
					495					500					505	
ATG	AAG	CTG	CTC	AAA	CTG	TAT	GCG	TGG	GAG	AGC	ATC	TTC	TGC	TCC	AGG	1587
Met	Lys	Leu	Leu	Lys	Leu	Tyr	Ala	Trp	Glu	Ser	Ile	Phe	Cys	Ser	Arg	
				510				515						520		
GTG	GAG	GTG	ACT	CGC	AGG	AAG	GAG	ATG	ACC	AGC	CTG	AGG	GCG	TTT	GCT	1635
Val	Glu	Val	Thr	Arg	Arg	Lys	Glu	Met	Thr	Ser	Leu	Arg	Ala	Phe	Ala	
			525					530					535			
GTC	TAC	ACT	TCC	ATC	TCC	ATC	TTC	ATG	AAC	ACA	GCC	ATC	CCC	ATT	GCT	1683
Val	Tyr	Ser	Ile	Ser	Ile	Phe	Met	Asn	Thr	Ala	Ile	Pro	Ile	Ala		
		540					545					550				
GCC	GTG	CTC	ATC	ACC	TTC	GTG	GGC	CAC	GTG	AGC	TTC	TTC	AAA	GAG	TCG	1731
Ala	Val	Leu	Ile	Thr	Phe	Val	Gly	His	Val	Ser	Phe	Phe	Lys	Glu	Ser	
		555				560					565					
GAC	TTG	TCA	CCC	TCG	GTG	GCC	TTT	GCC	TCC	CTC	TCT	CTC	TTC	CAC	ATC	1779
Asp	Leu	Ser	Pro	Ser	Val	Ala	Phe	Ala	Ser	Leu	Ser	Leu	Phe	His	Ile	
					575					580					585	
CTG	GTG	ACT	CCA	CTG	TTC	CTG	CTG	TCT	AGC	GTG	GTT	CGG	TCC	ACT	GTG	1827
Leu	Val	Thr	Pro	Leu	Phe	Leu	Leu	Ser	Ser	Val	Val	Arg	Ser	Thr	Val	
				590					595					600		
AAA	GCC	CTG	GTG	AGC	GTG	CAA	AAA	CTG	AGC	GAG	TTC	CTG	TCT	AGT	GCA	1875
Lys	Ala	Leu	Val	Ser	Val	Gln	Lys	Leu	Ser	Glu	Phe	Leu	Ser	Ser	Ala	
			605					610					615			
GAG	ATC	CGT	GAG	GAG	CAG	TGT	GCC	CCC	CGA	GAG	CCT	GCA	CCC	CAA	GGC	1923
Glu	Ile	Arg	Glu	Glu	Gln	Cys	Ala	Pro	Arg	Glu	Pro	Ala	Pro	Gln	Gly	
		620					625					630				
CAA	GCC	GGC	AAG	TAC	CAG	GCA	GTG	CCC	CTC	AAG	GTT	GTG	AAC	CGC	AAA	1971
Gln	Ala	Gly	Lys	Tyr	Gln	Ala	Val	Pro	Leu	Lys	Val	Val	Asn	Arg	Lys	
		635				640					645					
CGC	CCA	GCC	CGG	GAA	GAG	GTG	CGG	GAC	CTC	CTG	GGC	CCA	CTG	CAG	AGG	2019
Arg	Pro	Ala	Arg	Glu	Glu	Val	Arg	Asp	Leu	Leu	Gly	Pro	Leu	Gln	Arg	
		650			655					660					665	
CTG	GCC	CCT	AGC	ATG	GAC	GGG	GAT	GCT	GAC	AAC	TTC	TGT	GTG	CAG	ATC	2067
Leu	Ala	Pro	Ser	Met	Asp	Gly	Asp	Ala	Asp	Asn	Phe	Cys	Val	Gln	Ile	
				670				675						680		
ATC	GGA	GGC	TTC	TTC	ACC	TGG	ACC	CCT	GAT	GGA	ATC	CCC	ACT	CTG	TCC	2115
Ile	Gly	Gly	Phe	Phe	Thr	Trp	Thr	Pro	Asp	Gly	Ile	Pro	Thr	Leu	Ser	
			685					690					695			
AAC	ATC	ACC	ATC	CGT	ATT	CCC	CGA	GGT	CAG	CTA	ACC	ATG	ATT	GTG	GGG	2163
Asn	Ile	Thr	Ile	Arg	Ile	Pro	Arg	Gly	Gln	Leu	Thr	Met	Ile	Val	Gly	
		700					705					710				
CAG	GTG	GGC	TGC	GGC	AAG	TCC	TCG	CTC	CTC	CTC	GCC	ACC	CTG	GGG	GAG	2211
Gln	Val	Gly	Cys	Gly	Lys	Ser	Ser	Leu	Leu	Leu	Ala	Thr	Leu	Gly	Glu	
		715					720					725				

SUBSTITUTE SHEET (RULE 26)

110

ATG Met 730	CAG Gln 730	AAG Lys 730	GTG Val 730	TCG Ser 735	GGG Gly 735	GCC Ala 735	GTC Val 735	TTC Phe 735	TGG Trp 740	AAC Asn 740	AGC Ser 740	AAC Asn 740	CTT Leu 745	CCG Pro 745	GAC Asp 745	2259
AGC Ser 750	GAG Glu 750	GGG Gly 750	AGA Arg 750	GGA Gly 750	CCC Pro 755	CAG Gln 755	CAG Gln 755	CCC Pro 755	AGA Arg 755	GCG Ala 755	GGA Gly 755	GAC Asp 760	AGC Ser 760	AGC Ser 760	TGG Trp 760	2307
CTC Leu 765	GGA Gly 765	TAT Tyr 765	CAG Gln 765	GAG Glu 765	CAG Gln 765	AGG Arg 770	CCC Pro 770	CGT Arg 770	GGC Gly 770	TAC Tyr 775	GCA Ala 775	TCT Ser 775	CAG Gln 775	AAA Lys 775	CCA Pro 775	2355
TGG Trp 780	CTG Leu 780	CTA Leu 780	AAC Asn 780	GCC Ala 785	ACC Thr 785	GTG Val 785	GAA Glu 785	GAG Glu 785	AAC Asn 790	ATC Ile 790	ACC Thr 790	TTC Phe 790	GAG Glu 790	AGT Ser 790	CCC Pro 790	2403
TTC Phe 795	AAT Asn 795	CCG Pro 795	CAG Gln 795	CGG Arg 795	TAC Tyr 800	AAG Lys 800	ATG Met 800	GTC Val 800	ATC Ile 805	GAA Glu 805	GCC Ala 805	TGC Cys 805	TCC Ser 805	CTG Leu 805	CAG Gln 805	2451
CCG Pro 810	GAC Asp 810	ATA Ile 810	GAC Asp 810	ATC Ile 815	CTG Leu 815	CCC Pro 815	CAC His 815	GGA Gly 820	GAC Asp 820	CAG Gln 820	ACT Thr 820	CAG Gln 820	ATT Ile 820	GGG Gly 825	GAA Glu 825	2499
CGG Arg 830	GGC Gly 830	ATC Ile 830	AAC Asn 830	CTG Leu 830	TCT Ser 830	GGT Gly 835	GGT Gly 835	CAG Gln 835	CGT Arg 835	CCA Pro 835	GAT Asp 835	CAG Gln 840	TGT Cys 840	GGT Gly 840	CCA Pro 840	2547
GAG Glu 845	CCC Pro 845	TCT Ser 845	ACC Thr 845	AGC Ser 845	AGA Arg 845	CCA Pro 850	ATG Met 850	TTC Phe 850	GTC Val 850	TTC Phe 855	TTG Leu 855	GAT Asp 855	GAC Asp 855	CCC Pro 855	TTC Phe 855	2595
TCA Ser 860	GCT Ala 860	TTG Leu 860	GAT Asp 860	GTC Val 865	CAT His 865	CTG Leu 865	AGT Ser 865	GAC Asp 865	CAC His 870	CTG Leu 870	ATG Met 870	CAG Gln 870	GCC Ala 870	GGC Gly 870	ATC Ile 870	2643
CTT Leu 875	GAG Glu 875	CTG Leu 875	CTC Leu 875	CGG Arg 880	GAT Asp 880	GAC Asp 880	AAG Lys 880	AGG Arg 880	ACA Thr 885	GTG Val 885	GTC Val 885	TTG Leu 885	GTG Val 885	ACC Thr 885	CAC His 885	2691
AAG Lys 890	CTA Leu 890	CAG Gln 890	TAT Tyr 890	CTG Leu 895	CCT Pro 895	CAT His 895	GCA Ala 895	GAC Asp 900	TGG Trp 900	ATC Ile 900	ATT Ile 900	GCC Ala 900	ATG Met 900	AAG Lys 905	GAT Asp 905	2739
GGG Gly 910	ACC Thr 910	ATT Ile 910	CAG Gln 910	AGG Arg 910	GAA Glu 915	GGG Gly 915	ACG Thr 915	CTC Leu 915	AAG Lys 915	GAC Asp 915	TTC Phe 920	CAG Gln 920	AGG Arg 920	TCC Ser 920	GAG Glu 920	2787
TGC Cys 925	CAG Gln 925	CTC Leu 925	TTT Phe 925	GAG Glu 925	CAC His 930	TGG Trp 930	AAG Lys 930	ACC Thr 930	CTC Leu 930	ATG Met 930	AAC Asn 935	CGG Arg 935	CAG Gln 935	GAC Asp 935	CAA Gln 935	2835
GAG Glu 940	CTG Leu 940	GAG Glu 940	AAG Lys 940	GAG Glu 945	ACA Thr 945	GTC Val 945	ATG Met 945	GAG Glu 945	AGG Arg 945	AAA Lys 950	GCC Ala 950	TCA Ser 950	GAG Glu 950	CCA Pro 950	TCT Ser 950	2883
CAG Gln 955	GGC Gly 955	CTG Leu 955	CCC Pro 955	CGT Arg 960	GCC Ala 960	ATG Met 960	TCC Ser 960	TCC Ser 960	AGA Arg 965	GAC Asp 965	GGC Gly 965	CTT Leu 965	CTG Leu 965	CTG Leu 965	GAT Asp 965	2931
GAG Glu 970	GAA Glu 970	GAG Glu 970	GAG Glu 970	GAA Glu 975	GAG Glu 975	GAG Glu 975	GCA Ala 975	GCC Ala 980	GAA Glu 980	AGC Ser 980	GAG Glu 980	GAA Glu 980	GAT Asp 985	GAC Asp 985	AAC Asn 985	2979
TTA Leu 990	TCT Ser 990	TCA Ser 990	GTG Val 990	CTG Leu 990	CAT His 990	CAG Gln 990	CGA Arg 995	GCT Ala 995	AAG Lys 995	ATC Ile 995	CCC Pro 1000	TGG Trp 1000	CGA Arg 1000	TCC Ala 1000	TGC Cys 1000	3027

SUBSTITUTE SHEET (RULE 26)

111

ACT AAG TAT CTG TCC TCT GCT GGC ATT CTG CTC CTG TCC CTG CTT GTC Thr Lys Tyr Leu Ser Ser Ala Gly Ile Leu Leu Leu Ser Leu Leu Val 1005 1010 1015	3075
TTC TCC CAG CTG CTC AAG CAC ATG GTC TTG GTG GCC ATT GAT TAT TGG Phe Ser Gln Leu Leu Lys His Met Val Leu Val Ala Ile Asp Tyr Trp 1020 1025 1030	3123
CTG GCC AAG TGG ACG GAC AGT GCC CTG GTC CTG AGC CCC GCT GCC AGG Leu Ala Lys Trp Thr Asp Ser Ala Leu Val Leu Ser Pro Ala Ala Arg 1035 1040 1045	3171
AAC TGT TCG CTC AGC CAG GAA TGT GAC CTG GAC CAG TCT GTC TAT GCC Asn Cys Ser Leu Ser Gln Glu Cys Asp Leu Asp Gln Ser Val Tyr Ala 1050 1055 1060 1065	3219
ATG GTA TTC ACC TTG CTC TGC AGC CTG GGT ATC GTG CTG TGC CTG GTC Met Val Phe Thr Leu Leu Cys Ser Leu Gly Ile Val Leu Cys Leu Val 1070 1075 1080	3267
ACC TCT GTC ACT GTG GAG TGG ACG GGA CTG AAG GTG GCC AAG AGG CTA Thr Ser Val Thr Val Glu Trp Thr Gly Leu Lys Val Ala Lys Arg Leu 1085 1090 1095	3315
CAC CGC AGC CTG CTC AAC CGC ATC ATC CTG GCC CCC ATG AGG TTC TTT His Arg Ser Leu Leu Asn Arg Ile Ile Leu Ala Pro Met Arg Phe Phe 1100 1105 1110	3363
GAG ACC ACA CCC CTC GGG AGT ATC CTG AAC AGA TTT TCA TCC GAC TGT Glu Thr Thr Pro Leu Gly Ser Ile Leu Asn Arg Phe Ser Ser Asp Cys 1115 1120 1125	3411
AAC ACC ATT GAC CAG CAC ATC CCA TCC ACG CTG GAG TGT CTG AGC CGG Asn Thr Ile Asp Gln His Ile Pro Ser Thr Leu Glu Cys Leu Ser Arg 1130 1135 1140 1145	3459
TCC ACC CTG CTG TGT GTC TCC GCC CTG ACT GTC ATC TCC TAT GTC ACA Ser Thr Leu Leu Cys Val Ser Ala Leu Thr Val Ile Ser Tyr Val Thr 1150 1155 1160	3507
CCC GTG TTC CTC GTG GCC CTC TTA CCC CTA GCT GTT GTG TGC TAC TTC Pro Val Phe Leu Val Ala Leu Leu Pro Leu Ala Val Val Cys Tyr Phe 1165 1170 1175	3555
ATT CAG AAG TAC TTC CGA GTG GCA TCC AGG GAC CTG CAG CAG CTG GAC Ile Gln Lys Tyr Phe Arg Val Ala Ser Arg Asp Leu Gln Gln Leu Asp 1180 1185 1190	3603
GAC ACG ACG CAG CTC CCG CTC GTC TCA CAC TTT GCT GAA ACT GTG GAG Asp Thr Thr Gln Leu Pro Leu Val Ser His Phe Ala Glu Thr Val Glu 1195 1200 1205	3651
GGA CTC ACC ACC ATC CGT GCC TTC AGG TAC GAG GCC CGG TTC CAG CAG Gly Leu Thr Thr Ile Arg Ala Phe Arg Tyr Glu Ala Arg Phe Gln Gln 1210 1215 1220 1225	3699
AAG CTT CTA GAA TAT ACC GAC TCC AAC AAC ATC GCC TCC CTC TTC CTC Lys Leu Leu Glu Tyr Thr Asp Ser Asn Ile Ala Ser Leu Phe Leu 1230 1235 1240	3747
ACG GCA GCC AAC AGA TGG CTG GAA GTC TGC ATG GAG TAC ATC GGA GCG Thr Ala Ala Asn Arg Trp Leu Glu Val Cys Met Glu Tyr Ile Gly Ala 1245 1250 1255	3795
TGC GTG GTA CTC ATT GCG GCT GCC ACC TCC ATC TCC AAC TCC CTG CAC Cys Val Val Leu Ile Ala Ala Ala Thr Ser Ile Ser Asn Ser Leu His 1260 1265 1270	3843

SUBSTITUTE SHEET (RULE 26)

112

AGG GAA CTT TCT GCT GGC CTG GTG GGC CTG GGC CTC ACC TAT GCC TTG Arg Glu Leu Ser Ala Gly Leu Val Gly Leu Gly Leu Thr Tyr Ala Leu 1275 1280 1285	3891
ATG GTC TCC AAC TAC CTC AAC TGG ATG GTG AGG AAC CTG GCG GAC ATG Met Val Ser Asn Tyr Leu Asn Trp Met Val Arg Asn Leu Ala Asp Met 1290 1295 1300 1305	3939
GAG ATC CAG CTG GGG GCT GTG AAG AGG ATC CAC GCA CTC CTG AAA ACC Glu Ile Gln Leu Gly Ala Val Lys Arg Ile His Ala Leu Leu Lys Thr 1310 1315 1320	3987
GAG GCG GAG AGC TAT GAG GGG CTC CTG GCG CCG TCG TTG ATC CCC AAG Glu Ala Glu Ser Tyr Glu Gly Leu Ala Pro Ser Leu Ile Pro Lys 1325 1330 1335	4035
AAC TGG CCA GAC CAA GGG AAG ATC CAA ATT CAG AAC CTG AGC GTG CGC Asn Trp Pro Asp Gln Gly Lys Ile Gln Ile Gln Asn Leu Ser Val Arg 1340 1345 1350	4083
TAT GAC AGC TCC CTG AAG CCA GTG CTG AAG CAT GTC AAC ACC CTC ATC Tyr Asp Ser Ser Leu Lys Pro Val Leu Lys His Val Asn Thr Leu Ile 1355 1360 1365	4131
TCC CCG GGG CAG AAG ATC GGG ATC TGC GGC CGC ACA GGC AGC GGG AAG Ser Pro Gly Gln Lys Ile Gly Ile Cys Gly Arg Thr Gly Ser Gly Lys 1370 1375 1380 1385	4179
TCC TCC TTC TCC CTG GCC TTT TTC CGA ATG GTG GAC ATG TTT GAA GGA Ser Ser Phe Ser Leu Ala Phe Phe Arg Met Val Asp Met Phe Glu Gly 1390 1395 1400	4227
CGC ATC ATC ATT GAT GGC ATC GAC ATC GCC AAG CTG CCA CTT CAC ACG Arg Ile Ile Ile Asp Gly Ile Asp Ile Ala Lys Leu Pro Leu His Thr 1405 1410 1415	4275
CTG CGC TCA CGC CTG TCC ATC ATC CTA CAG GAC CCC GTC CTC TTC AGC Leu Arg Ser Arg Leu Ser Ile Ile Leu Gln Asp Pro Val Leu Phe Ser 1420 1425 1430	4323
GGC ACG ATC AGA TTC AAC CTG GAC CCC GAG AAG AAA TGC TCA GAC AGC Gly Thr Ile Arg Phe Asn Leu Asp Pro Glu Lys Lys Cys Ser Asp Ser 1435 1440 1445	4371
ACA CTG TGG GAG GCC CTG GAG ATC GCC CAG CTG AAG CTG GTA GTG AAG Thr Leu Trp Glu Ala Leu Glu Ile Ala Gln Leu Lys Leu Val Val Lys 1450 1455 1460 1465	4419
GCA CTG CCA GGA GGC CTA GAT GCC ATC ATC ACA GAA GGA GGG GAG AAT Ala Leu Pro Gly Gly Leu Asp Ala Ile Ile Thr Glu Gly Gly Glu Asn 1470 1475 1480	4467
TTT AGC CAG GGC CAG AGG CAG CTG TTC TGC CTG GCC CGG GCC TTC GTG Phe Ser Gln Gly Gln Arg Gln Leu Phe Cys Leu Ala Arg Ala Phe Val 1485 1490 1495	4515
AGG AAG ACC AGC ATC TTC ATC ATG GAT GAA GCA ACC GCC TCC ATC GAC Arg Lys Thr Ser Ile Phe Ile Met Asp Glu Ala Thr Ala Ser Ile Asp 1500 1505 1510	4563
ATG GCT ACG GAG AAC ATC CTC CAG AAG GTG GTG ATG ACA GCC TTC GCA Met Ala Thr Glu Asn Ile Leu Gln Lys Val Val Met Thr Ala Phe Ala 1515 1520 1525	4611
GAC CGC ACG GTG GTC ACC ATC GCG CAT CGT GTG CAC ACC ATC CTG AGT Asp Arg Thr Val Val Thr Ile Ala His Arg Val His Thr Ile Leu Ser 1530 1535 1540 1545	4659

SUBSTITUTE SHEET (RULE 26)

113

GCA GAC CTG GTG ATG GTC CTC AAG AGG GGT GCT ATC CTG GAG TTT GAC 4707  
 Ala Asp Leu Val Met Val Leu Lys Arg Gly Ala Ile Leu Glu Phe Asp  
 1550 1555 1560

AAG CCA GAG ACG CTC CTC AGC CAG AAG GAC AGC GTG TTC GCC TCC TTT 4755  
 Lys Pro Glu Thr Leu Leu Ser Gln Lys Asp Ser Val Phe Ala Ser Phe  
 1565 1570 1575

GTC CGT GCG GAC AAG TGA CTTACCG GAGCCAAAGT GCCACCCCGC GCCTCGCTTG 4810  
 Val Arg Ala Asp Lys  
 1580

CTTGCCTAGG ATTTCTAACT GCAAATCACT TGTAATAAAA TTAATTCTTT GCTAAAAAAA 4870  
 AAAAAAA 4877

(36) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

Met Pro Leu Ala Phe Cys Gly Thr Glu Asn His Ser Ala Ala Tyr Arg  
 1 5 10 15

Val Asp Gln Gly Val Leu Asn Asn Gly Cys Phe Val Asp Ala Leu Asn  
 20 25 30

Val Val Pro His Val Phe Leu Leu Phe Ile Thr Phe Pro Ile Leu Phe  
 35 40 45

Ile Gly Trp Gly Ser Gln Ser Ser Lys Val His Ile His His Ser Thr  
 50 55 60

Trp Leu His Phe Pro Gly His Asn Leu Arg Trp Ile Leu Thr Phe Ile  
 65 70 75 80

Leu Leu Phe Val Leu Val Cys Glu Ile Ala Glu Gly Ile Leu Ser Asp  
 85 90 95

Gly Val Thr Glu Ser Arg His Leu His Leu Tyr Met Pro Ala Gly Met  
 100 105 110

Ala Phe Met Ala Ala Ile Thr Ser Val Val Tyr Tyr His Asn Ile Glu  
 115 120 125

Thr Ser Asn Phe Pro Lys Leu Leu Ile Ala Leu Leu Ile Tyr Trp Thr  
 130 135 140

Leu Ala Phe Ile Thr Lys Thr Ile Lys Phe Val Lys Phe Tyr Asp His  
 145 150 155 160

Ala Ile Gly Phe Ser Gln Leu Arg Phe Cys Leu Thr Gly Leu Leu Val  
 165 170 175

Ile Leu Tyr Gly Met Leu Leu Leu Val Glu Val Asn Val Ile Arg Val  
 180 185 190

Arg Arg Tyr Ile Phe Phe Lys Thr Pro Arg Glu Val Lys Pro Pro Glu  
 195 200 205

SUBSTITUTE SHEET (RULE 26)



114

Asp Leu Gln Asp Leu Gly Val Arg Phe Leu Gln Pro Phe Val Asn Leu  
 210 215 220  
 Leu Ser Lys Gly Thr Tyr Trp Trp Met Asn Ala Phe Ile Lys Thr Ala  
 225 230 235 240  
 His Lys Lys Pro Ile Asp Leu Arg Ala Ile Ala Lys Leu Pro Ile Ala  
 245 250 255  
 Met Arg Ala Leu Thr Asn Tyr Gln Arg Leu Cys Val Ala Phe Asp Ala  
 260 265 270  
 Gln Ala Arg Lys Asp Thr Gln Ser Pro Gln Gly Ala Arg Ala Ile Trp  
 275 280 285  
 Arg Ala Leu Cys His Ala Phe Gly Arg Arg Leu Ile Leu Ser Ser Thr  
 290 295 300  
 Phe Arg Ile Leu Ala Asp Leu Leu Gly Phe Ala Gly Pro Leu Cys Ile  
 305 310 315 320  
 Phe Gly Ile Val Asp His Leu Gly Lys Glu Asn His Val Phe Gln Pro  
 325 330 335  
 Lys Thr Gln Phe Leu Gly Val Tyr Phe Val Ser Ser Gln Glu Phe Leu  
 340 345 350  
 Gly Asn Ala Tyr Val Leu Ala Val Leu Leu Phe Leu Ala Leu Leu Leu  
 355 360 365  
 Gln Arg Thr Phe Leu Gln Ala Ser Tyr Tyr Val Ala Ile Glu Thr Gly  
 370 375 380  
 Ile Asn Leu Arg Gly Ala Ile Gln Thr Lys Ile Tyr Asn Lys Ile Met  
 385 390 395 400  
 His Met Ser Thr Ser Asn Leu Ser Met Gly Glu Met Thr Ala Gly Gln  
 405 410 415  
 Ile Cys Asn Leu Val Ala Ile Asp Thr Asn Gln Leu Met Trp Phe Phe  
 420 425 430  
 Phe Leu Cys Pro Asn Leu Trp Thr Met Pro Val Gln Ile Ile Val Gly  
 435 440 445  
 Val Ile Leu Leu Tyr Tyr Ile Leu Gly Val Ser Ala Leu Ile Gly Ala  
 450 455 460  
 Ala Val Ile Ile Leu Leu Ala Pro Val Gln Tyr Phe Val Ala Thr Lys  
 465 470 475 480  
 Leu Ser Gln Ala Gln Arg Thr Thr Leu Glu His Ser Asn Glu Arg Leu  
 485 490 495  
 Lys Gln Thr Asn Glu Met Leu Arg Gly Met Lys Leu Leu Lys Leu Tyr  
 500 505 510  
 Ala Trp Glu Ser Ile Phe Cys Ser Arg Val Glu Val Thr Arg Arg Lys  
 515 520 525  
 Glu Met Thr Ser Leu Arg Ala Phe Ala Val Tyr Thr Ser Ile Ser Ile  
 530 535 540  
 Phe Met Asn Thr Ala Ile Pro Ile Ala Ala Val Leu Ile Thr Phe Val  
 545 550 555 560  
 Gly His Val Ser Phe Phe Lys Glu Ser Asp Leu Ser Pro Ser Val Ala

SUBSTITUTE SHEET (RULE 26)

115

565										570					575				
Phe	Ala	Ser	Leu	Ser	Leu	Phe	His	Ile	Leu	Val	Thr	Pro	Leu	Phe	Leu				
			580					585					590						
Leu	Ser	Ser	Val	Val	Arg	Ser	Thr	Val	Lys	Ala	Leu	Val	Ser	Val	Gln				
		595					600					605							
Lys	Leu	Ser	Glu	Phe	Leu	Ser	Ser	Ala	Glu	Ile	Arg	Glu	Glu	Gln	Cys				
	610					615					620								
Ala	Pro	Arg	Glu	Pro	Ala	Pro	Gln	Gly	Gln	Ala	Gly	Lys	Tyr	Gln	Ala				
625					630					635					640				
Val	Pro	Leu	Lys	Val	Val	Asn	Arg	Lys	Arg	Pro	Ala	Arg	Glu	Glu	Val				
				645					650					655					
Arg	Asp	Leu	Leu	Gly	Pro	Leu	Gln	Arg	Leu	Ala	Pro	Ser	Met	Asp	Gly				
			660					665					670						
Asp	Ala	Asp	Asn	Phe	Cys	Val	Gln	Ile	Ile	Gly	Gly	Phe	Phe	Thr	Trp				
		675					680					685							
Thr	Pro	Asp	Gly	Ile	Pro	Thr	Leu	Ser	Asn	Ile	Thr	Ile	Arg	Ile	Pro				
		690					695					700							
Arg	Gly	Gln	Leu	Thr	Met	Ile	Val	Gly	Gln	Val	Gly	Cys	Gly	Lys	Ser				
705					710						715				720				
Ser	Leu	Leu	Leu	Ala	Thr	Leu	Gly	Glu	Met	Gln	Lys	Val	Ser	Gly	Ala				
				725					730					735					
Val	Phe	Trp	Asn	Ser	Asn	Leu	Pro	Asp	Ser	Glu	Gly	Arg	Gly	Pro	Gln				
			740					745					750						
Gln	Pro	Arg	Ala	Gly	Asp	Ser	Ser	Trp	Leu	Gly	Tyr	Gln	Glu	Gln	Arg				
		755					760					765							
Pro	Arg	Gly	Tyr	Ala	Ser	Gln	Lys	Pro	Trp	Leu	Leu	Asn	Ala	Thr	Val				
		770					775					780							
Glu	Glu	Asn	Ile	Thr	Phe	Glu	Ser	Pro	Phe	Asn	Pro	Gln	Arg	Tyr	Lys				
785					790						795				800				
Met	Val	Ile	Glu	Ala	Cys	Ser	Leu	Gln	Pro	Asp	Ile	Asp	Ile	Leu	Pro				
				805					810					815					
His	Gly	Asp	Gln	Thr	Gln	Ile	Gly	Glu	Arg	Gly	Ile	Asn	Leu	Ser	Gly				
			820					825					830						
Gly	Gln	Arg	Pro	Asp	Gln	Cys	Gly	Pro	Glu	Pro	Ser	Thr	Ser	Arg	Pro				
		835					840					845							
Met	Phe	Val	Phe	Leu	Asp	Asp	Pro	Phe	Ser	Ala	Leu	Asp	Val	His	Leu				
		850				855					860								
Ser	Asp	His	Leu	Met	Gln	Ala	Gly	Ile	Leu	Glu	Leu	Leu	Arg	Asp	Asp				
865					870						875				880				
Lys	Arg	Thr	Val	Val	Leu	Val	Thr	His	Lys	Leu	Gln	Tyr	Leu	Pro	His				
				885					890					895					
Ala	Asp	Trp	Ile	Ile	Ala	Met	Lys	Asp	Gly	Thr	Ile	Gln	Arg	Glu	Gly				
			900					905					910						
Thr	Leu	Lys	Asp	Phe	Gln	Arg	Ser	Glu	Cys	Gln	Leu	Phe	Glu	His	Trp				
		915					920					925							

SUBSTITUTE SHEET (RULE 26)

116

Lys Thr Leu Met Asn Arg Gln Asp Gln Glu Leu Glu Lys Glu Thr Val  
 930 935 940  
 Met Glu Arg Lys Ala Ser Glu Pro Ser Gln Gly Leu Pro Arg Ala Met  
 945 950 955 960  
 Ser Ser Arg Asp Gly Leu Leu Leu Asp Glu Glu Glu Glu Glu Glu  
 965 970 975  
 Ala Ala Glu Ser Glu Glu Asp Asp Asn Leu Ser Ser Val Leu His Gln  
 980 985 990  
 Arg Ala Lys Ile Pro Trp Arg Ala Cys Thr Lys Tyr Leu Ser Ser Ala  
 995 1000 1005  
 Gly Ile Leu Leu Leu Ser Leu Leu Val Phe Ser Gln Leu Leu Lys His  
 1010 1015 1020  
 Met Val Leu Val Ala Ile Asp Tyr Trp Leu Ala Lys Trp Thr Asp Ser  
 1025 1030 1035 1040  
 Ala Leu Val Leu Ser Pro Ala Ala Arg Asn Cys Ser Leu Ser Gln Glu  
 1045 1050 1055  
 Cys Asp Leu Asp Gln Ser Val Tyr Ala Met Val Phe Thr Leu Leu Cys  
 1060 1065 1070  
 Ser Leu Gly Ile Val Leu Cys Leu Val Thr Ser Val Thr Val Glu Trp  
 1075 1080 1085  
 Thr Gly Leu Lys Val Ala Lys Arg Leu His Arg Ser Leu Leu Asn Arg  
 1090 1095 1100  
 Ile Ile Leu Ala Pro Met Arg Phe Phe Glu Thr Thr Pro Leu Gly Ser  
 1105 1110 1115 1120  
 Ile Leu Asn Arg Phe Ser Ser Asp Cys Asn Thr Ile Asp Gln His Ile  
 1125 1130 1135  
 Pro Ser Thr Leu Glu Cys Leu Ser Arg Ser Thr Leu Leu Cys Val Ser  
 1140 1145 1150  
 Ala Leu Thr Val Ile Ser Tyr Val Thr Pro Val Phe Leu Val Ala Leu  
 1155 1160 1165  
 Leu Pro Leu Ala Val Val Cys Tyr Phe Ile Gln Lys Tyr Phe Arg Val  
 1170 1175 1180  
 Ala Ser Arg Asp Leu Gln Gln Leu Asp Asp Thr Thr Gln Leu Pro Leu  
 1185 1190 1195 1200  
 Val Ser His Phe Ala Glu Thr Val Glu Gly Leu Thr Thr Ile Arg Ala  
 1205 1210 1215  
 Phe Arg Tyr Glu Ala Arg Phe Gln Gln Lys Leu Leu Glu Tyr Thr Asp  
 1220 1225 1230  
 Ser Asn Asn Ile Ala Ser Leu Phe Leu Thr Ala Ala Asn Arg Trp Leu  
 1235 1240 1245  
 Glu Val Cys Met Glu Tyr Ile Gly Ala Cys Val Val Leu Ile Ala Ala  
 1250 1255 1260  
 Ala Thr Ser Ile Ser Asn Ser Leu His Arg Glu Leu Ser Ala Gly Leu  
 1265 1270 1275 1280  
 Val Gly Leu Gly Leu Thr Tyr Ala Leu Met Val Ser Asn Tyr Leu Asn

SUBSTITUTE SHEET (RULE 26)

117

1285										1290					1295				
Trp	Met	Val	Arg	Asn	Leu	Ala	Asp	Met	Glu	Ile	Gln	Leu	Gly	Ala	Val				
			1300						1305				1310						
Lys	Arg	Ile	His	Ala	Leu	Leu	Lys	Thr	Glu	Ala	Glu	Ser	Tyr	Glu	Gly				
		1315					1320					1325							
Leu	Leu	Ala	Pro	Ser	Leu	Ile	Pro	Lys	Asn	Trp	Pro	Asp	Gln	Gly	Lys				
		1330				1335					1340								
Ile	Gln	Ile	Gln	Asn	Leu	Ser	Val	Arg	Tyr	Asp	Ser	Ser	Leu	Lys	Pro				
1345					1350				1355						1360				
Val	Leu	Lys	His	Val	Asn	Thr	Leu	Ile	Ser	Pro	Gly	Gln	Lys	Ile	Gly				
			1365						1370					1375					
Ile	Cys	Gly	Arg	Thr	Gly	Ser	Gly	Lys	Ser	Ser	Phe	Ser	Leu	Ala	Phe				
			1380					1385					1390						
Phe	Arg	Met	Val	Asp	Met	Phe	Glu	Gly	Arg	Ile	Ile	Ile	Asp	Gly	Ile				
		1395					1400						1405						
Asp	Ile	Ala	Lys	Leu	Pro	Leu	His	Thr	Leu	Arg	Ser	Arg	Leu	Ser	Ile				
		1410				1415					1420								
Ile	Leu	Gln	Asp	Pro	Val	Leu	Phe	Ser	Gly	Thr	Ile	Arg	Phe	Asn	Leu				
1425					1430				1435						1440				
Asp	Pro	Glu	Lys	Lys	Cys	Ser	Asp	Ser	Thr	Leu	Trp	Glu	Ala	Leu	Glu				
			1445						1450					1455					
Ile	Ala	Gln	Leu	Lys	Leu	Val	Val	Lys	Ala	Leu	Pro	Gly	Gly	Leu	Asp				
		1460						1465					1470						
Ala	Ile	Ile	Thr	Glu	Gly	Gly	Glu	Asn	Phe	Ser	Gln	Gly	Gln	Arg	Gln				
		1475					1480					1485							
Leu	Phe	Cys	Leu	Ala	Arg	Ala	Phe	Val	Arg	Lys	Thr	Ser	Ile	Phe	Ile				
		1490				1495					1500								
Met	Asp	Glu	Ala	Thr	Ala	Ser	Ile	Asp	Met	Ala	Thr	Glu	Asn	Ile	Leu				
1505					1510				1515						1520				
Gln	Lys	Val	Val	Met	Thr	Ala	Phe	Ala	Asp	Arg	Thr	Val	Val	Thr	Ile				
				1525					1530					1535					
Ala	His	Arg	Val	His	Thr	Ile	Leu	Ser	Ala	Asp	Leu	Val	Met	Val	Leu				
			1540					1545					1550						
Lys	Arg	Gly	Ala	Ile	Leu	Glu	Phe	Asp	Lys	Pro	Glu	Thr	Leu	Leu	Ser				
		1555					1560					1565							
Gln	Lys	Asp	Ser	Val	Phe	Ala	Ser	Phe	Val	Arg	Ala	Asp	Lys						
		1570				1575						1580							

SUBSTITUTE SHEET (RULE 26)

- 118 -

**WHAT IS CLAIMED IS:**

1. A sulfonyleurea receptor encoded by the nucleic acid sequence of SEQ ID NO: 1.
2. The sulfonyleurea receptor of claim 1 selected from the group consisting of mouse, rat, and hamster sulfonyleurea receptor.
3. A sulfonyleurea receptor encoded by the nucleic acid sequence selected from the group consisting of SEQ ID NO: 4 and SEQ ID NO: 6.
4. The sulfonyleurea receptor of claim 3 encoded by the nucleic acid sequence of SEQ ID NO: 4.
5. The sulfonyleurea receptor of claim 4 further comprising SEQ ID NO: 2.
6. The sulfonyleurea receptor of claim 3 encoded by the nucleic acid sequence of SEQ ID NO: 6.
7. The sulfonyleurea receptor of claim 4 further comprising SEQ ID NO: 3.
8. A sulfonyleurea receptor encoded by the amino acid sequence selected from the group consisting of SEQ ID NO: 5 and SEQ ID NO: 7.
9. A purified nucleic acid sequence encoding a sulfonyleurea receptor selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 6.
10. A purified amino acid sequence encoding a sulfonyleurea receptor selected from the group consisting of SEQ ID NO: 5 and SEQ ID NO: 7.

- 119 -

11. An expression vector comprising a nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 6.
12. A host cell capable of expressing the nucleic acid sequence selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 4, and SEQ ID NO: 6.
13. A cell culture capable of expressing a sulfonyleurea receptor encoded by the amino acid sequence selected from the group consisting of SEQ ID NO: 5 and SEQ ID NO: 7.
14. A protein preparation comprising a sulfonyleurea receptor encoded by the amino acid sequence selected from the group consisting of SEQ ID NO: 5 and SEQ ID NO: 7.
15. A transgenic non-human mammal all of whose germ cells and somatic cells contain a recombinant sequence encoding a sulfonyleurea receptor introduced into said mammal, or an ancestor of said mammal.
16. The mammal of claim 15 wherein said sequence encoding a sulfonyleurea receptor is selected from the group consisting of SEQ ID NO: 1, 4, and 6.
17. A monoclonal antibody capable of binding to a sequence encoding a sulfonyleurea receptor.
18. The monoclonal antibody of claim 17 wherein said sequence encoding a sulfonyleurea receptor is selected from the group consisting of SEQ ID NO: 5 and 7.
19. A method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising

- 120 -

obtaining a sample comprising patient nucleic acids from a patient tissue sample;

amplifying sulfonylurea receptor specific nucleic acids from said patient nucleic acids to produce a test  
5 fragment;

obtaining a sample comprising control nucleic acids from a control tissue sample;

amplifying control nucleic acids encoding wild type sulfonylurea receptor to produce a control fragment;

10 comparing the test fragment with the control fragment to detect the presence of a sequence difference in the test fragment, wherein a difference in said test fragment indicates persistent hyperinsulinemic hypoglycemia of infancy.

15 20. The method of claim 19 wherein a sequence difference is selected from the group consisting of a nucleic acid transition and a restriction digest pattern alteration.

21. The method of claim 19 wherein a sequence difference is a nucleic acid transition.

20 22. The method of claim 21 wherein said nucleic acid transition is a G to A transition at nucleic acid position 750 of SEQ ID NO: 26.

23. The method of claim 21 wherein a nucleic acid transition results at position 27 of SEQ ID NO: 29.

25 24. The method of claim 19 wherein a sequence difference is a restriction digest pattern alteration.

25. The method of claim 24 wherein said restriction digest pattern alteration comprises test fragments of about 304 base pairs and about 123 base pairs  
30 upon digestion with MspI.

- 121 -

26. The method of claim 24 wherein said restriction digest pattern alteration comprises a test fragment of about 146 base pairs upon digestion with *NciI*.

27. The method of claim 19 wherein said  
5 amplification step comprises performing the polymerase chain reaction.

28. The method of claim 27 wherein the polymerase chain reaction comprises using a pair of primers, wherein one primer within said pair is selected from the group consisting  
10 of SEQ ID NOS: 16-24.

29. The method of claim 27 wherein said polymerase chain reaction comprises the use of two primers, a first primer selected from the group consisting of SEQUENCE ID NOS: 17, 18, 21, and 23, and a second primer selected from the  
15 group consisting of SEQUENCE ID NOS: 16, 19, 20, 22, and 24.

30. The method of claim 19 wherein said tissue sample is selected from the group consisting of pancreatic tissue, blood, serum, saliva, sputum, mucus, bone marrow, urine, lymph, and a tear.

20 31. A method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising  
obtaining a sample comprising patient mRNA from a patient tissue sample;  
reverse transcribing said mRNA into cDNA to produce  
25 patient cDNA;  
amplifying sulfonylurea receptor specific cDNA from said patient cDNA to produce a test fragment;  
obtaining a sample comprising control nucleic acids from a control tissue sample;  
30 amplifying control DNA encoding wild type sulfonylurea receptor to produce a control fragment;



- 122 -

digesting said test fragment and said control fragment with *MspI*;

comparing the test fragment with the control fragment to detect test fragments having about 304 base pairs and about 123 base pairs, wherein said test fragments indicate persistent hyperinsulinemic hypoglycemia of infancy.

32. A method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising

obtaining a sample comprising patient genomic DNA from a patient tissue sample;

amplifying sulfonylurea receptor specific DNA from said patient genomic DNA to produce a test fragment;

obtaining a sample comprising control nucleic acids from a control tissue sample;

amplifying control DNA encoding wild type sulfonylurea receptor to produce a control fragment;

comparing the test fragment with the control fragment to detect a test fragment having G to A transition at nucleic acid position 750 of SEQ ID NO: 26, wherein said test fragment indicates persistent hyperinsulinemic hypoglycemia of infancy.

33. A method of detecting persistent hyperinsulinemic hypoglycemia of infancy comprising

obtaining a sample comprising patient mRNA from a patient tissue sample;

reverse transcribing said mRNA into cDNA to produce patient cDNA;

amplifying sulfonylurea receptor specific cDNA from said patient cDNA to produce a test fragment;

obtaining a sample comprising control nucleic acids from a control tissue sample;

amplifying control DNA encoding wild type sulfonylurea receptor to produce a control fragment;

digesting said test fragment and said control fragment with *NciI*; and

- 123 -

comparing the test fragment with the control fragment to detect a test fragment of about 146 base pairs wherein said test fragment indicates persistent hyperinsulinemic hypoglycemia of infancy.

5           34. A method of claim 32, 33, or 34, wherein said amplification step comprises performing the polymerase chain reaction.

35. A sulfonylurea receptor encoded by the nucleic acid sequence of SEQ ID NO: 26.

10           36. The sulfonylurea receptor of claim 35 which is a human sulfonylurea receptor.

37. A sulfonylurea receptor encoded by the amino acid sequence of SEQ ID NO: 28.

15           38. A purified nucleic acid sequence encoding a sulfonylurea receptor comprising the sequence of SEQ ID NO: 26.

39. A purified amino acid sequence encoding a sulfonylurea receptor comprising the sequence of SEQ ID NO: 28.

20           40. An expression vector comprising a nucleic acid sequence of SEQ ID NO: 26.

41. A host cell capable of expressing the nucleic acid sequence of SEQ ID NO: 26.

25           42. A cell culture capable of expressing a sulfonylurea receptor encoded by the amino acid sequence of SEQ ID NO: 28.

- 124 -

43. A protein preparation comprising a sulfonylurea receptor encoded by the amino acid sequence of SEQ ID NO: 28.

44. A transgenic non-human mammal having all or  
5 part of its own sulfonylurea receptor sequence suppressed.

45. A transgenic non-human mammal comprising a recombinant wild type sequence for all or part of the sulfonylurea receptor sequence.

46. The transgenic non-human mammal of claim 45  
10 wherein said sulfonylurea receptor sequence is SEQ ID NO: 26.

47. A monoclonal antibody capable of binding to a sequence encoding a sulfonylurea receptor, wherein said sequence encoding a sulfonylurea receptor is SEQ ID NO: 28.

48. A diagnostic kit for detecting persistent  
15 hyperinsulinemic hypoglycemia of infancy comprising in one or more containers a pair of primers, wherein one primer within said pair is complementary to a region of the sulfonylurea receptor, wherein one of said pair of primers is selected from the group consisting of SEQ ID NOS: 16-24, a probe  
20 specific to the amplified product, and a means for visualizing amplified DNA, and optionally including one or more size markers, and positive and negative controls.

49. The diagnostic kit of claim 48 wherein said means for visualizing amplified DNA is selected from the  
25 group consisting of fluorescent stain, <sup>32</sup>P, and biotin.

50. The sequence encoded by a SEQ ID NO selected from the group consisting of SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

30

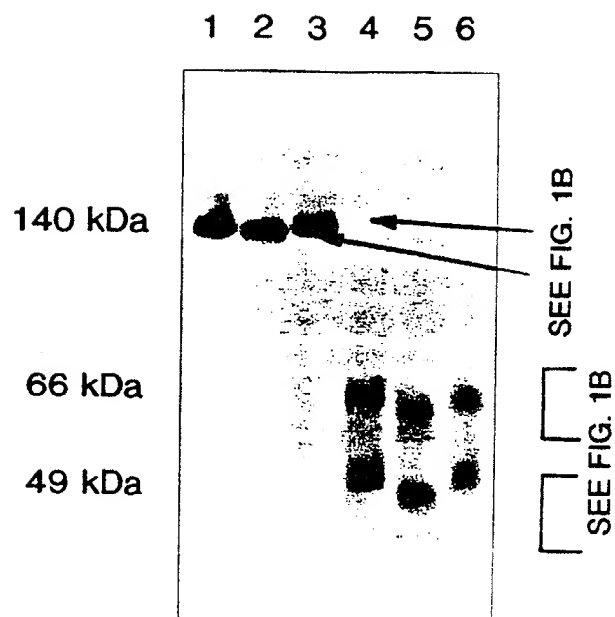


FIG. 1A

RECTIFIED SHEET (RULE 91)

[illegible]

**FIG. 1B**

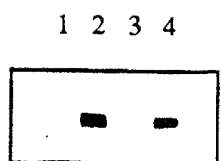


FIG. 2A

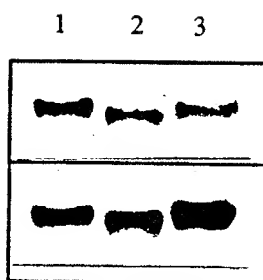


FIG. 2B

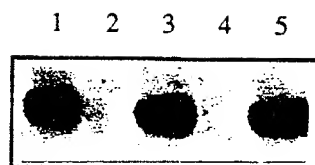


FIG. 2C

RECTIFIED SHEET (RULE 91)

Figure 3

Figure 3A



BNSDOCID: <WO\_\_9528411A1\_I\_>

```

1291
consensus .....VGI -GRIGSGKST LLLAF-RH-E --EGEIIIDG ---A-I-L-D -AKAF--IPQ DPVLFSGTFA QNLDPF--WS DEE-WK-LE- VGLK-VWE--
rat-sur .....MIGI cGRIGSGKSt FslAFFRMvD mEGcIIIDG ldiAkIpLht lRsrLsIIQ DPVLFSGTAR QNLDPekKCS DstlWaeLEl eqlKlVWkaL
dvhuar tIngGekVGI vGRIGSGKSt LtLqLFRInE saEGEIIIDG lniAkIqLhd lRfkitIIPQ DPVLFSGsLR mNLDPFsqYS DEEWtsLEl ehLkdfVsaL
lei-mdr qIePrekVGI vGRIGSGKST LLLtFMRHvE VcgGvIhVng remsaygLeD vRchfsmIPQ DPVLFdGTVR QNVDPFleas seEVWaeLEl VGLrerVase
Hu-CTTR SISPGqrVGI lGRIGSGKST LLsAFLRMin I.kGDleIDG VsmasVtLqE WRKAFgVItQ kvfIFSGTFA QNLDPngkWK DEEINWkaDe VGLKsVIEqF
Hu-CTTR SISPGqrVGI lGRIGSGKST LLsAFLRMin t.EGEIqIDG VsmasVtLqQ WRKAFgVItQ kvfIFSGTFA QNLDPyeqWS DqETWkaDe VGLsVIEqF

1385
consensus PG-LDF--VE GG-NLS-QQR QLMCLARALL -K-KI-I-LD EATAHIDP-T DQIIQRTV-- AFADCTVITI AHRI-TILD- -RVLVL-EG- V-EYDSPQ-L
rat-sur PGqLDelIE GGnFSqQQR QLFCLARAFv sKtsITf.MD EATAsIDeat EnIIQkvVat AFADrTVVTI AHRVhTILsa dIVMVLkrGe IIEFDkPekL
dvhuar PdkIDheceE GGnLSvQQR QLVCLARALL sKtKILV.LD EATAeVDleT DdIIQsTlct qFeDCTVITI AHRlnTIMDY tRVIvLDkGe IqEygePadL
lei-mdr seqIDsrviE GGnYSvQQR QLMCHARALL krgsqTILMD EATANIDPel DcqIQaTVma AFsaYTVITI AHRlhtVaQY dklIvMDhGv VaEmgSPreL
Hu-CTTR PGqLnFtIvD GGyvLSHGhk QLMCLARsvL sKaKILL.LD EpsAHIDPIT yQVlrRvIkq AFaGCTVilc eHRIeemLDC qRfLVIEEn VmQYDSIOaL
Hu-CTTR PGqLDsvIvD GGcvLSHGhk QLMCLARsvL sKaKILL.LD EpsAHIDPvT yQIIRtIkkq AFADCTVilc eHRIeemLEC qqfLVIEEnk VmQYDSIOkL

1485
consensus LS--S-FFOA IS-S--VR-F K.....
rat-sur LSqtdsvf.. ....sFVRed K.....
dvhuar LqqrglFYsa akDagLV... .....
lei-mdr vmmhqsmfhs mvESlgsRgs Kdfyellmqr
Hu-CTTR LSekSIFQOA ISaSEkmRFF qgRhSSKkKp
Hu-CTTR LnerSIFQOA ISpSDrVklF phRnSSKkKs

```

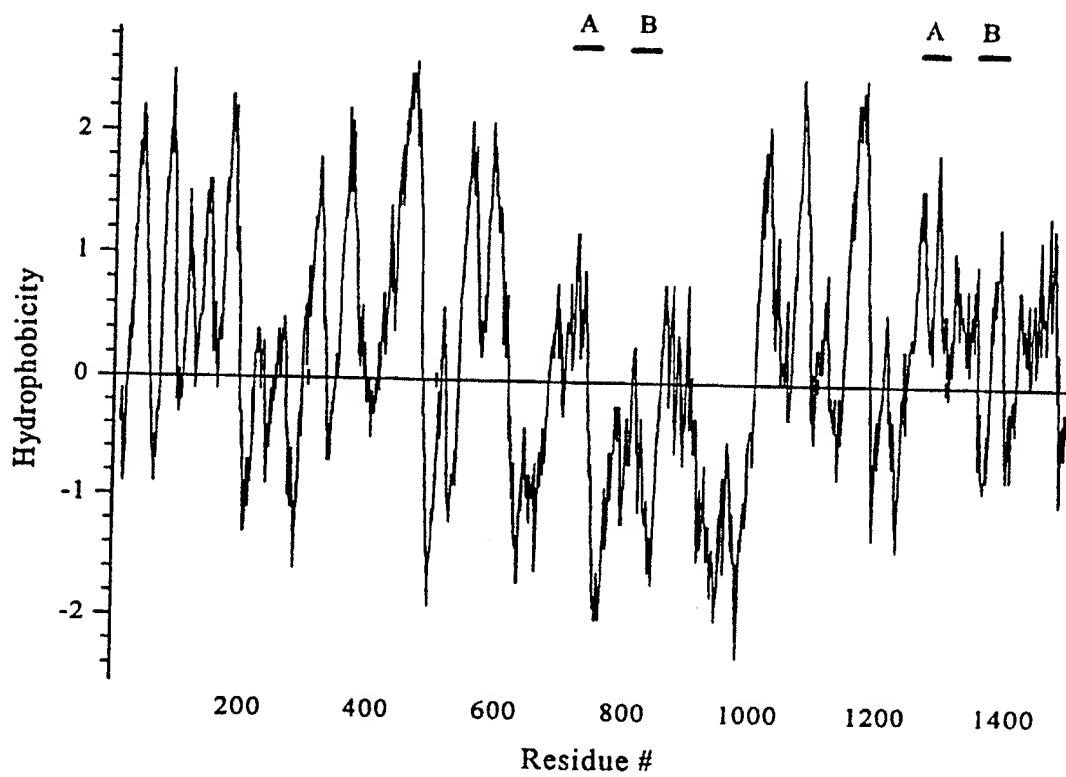
Figure 5

A B C D



FIG. 6

RECTIFIED SHEET (RULE 91)

**Figure 7**

RECTIFIED SHEET (RULE 91)

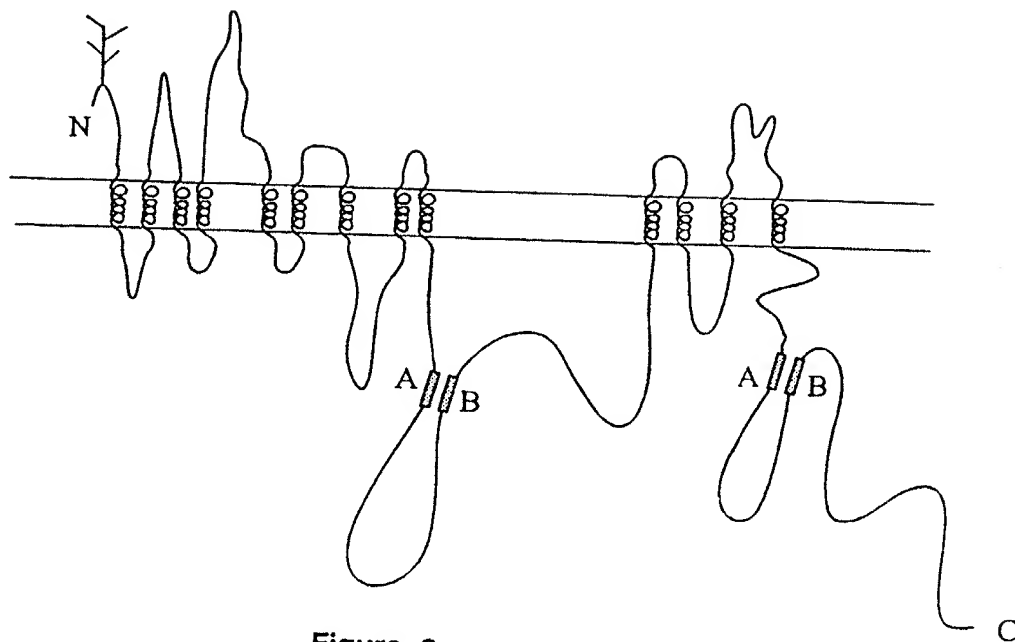


Figure 8

RECTIFIED SHEET (RULE 91)

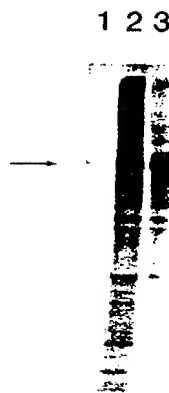


FIG. 9



FIG. 9A

RECTIFIED SHEET (RULE 91)

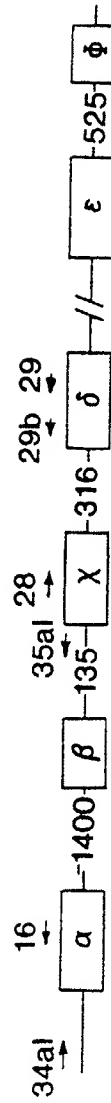
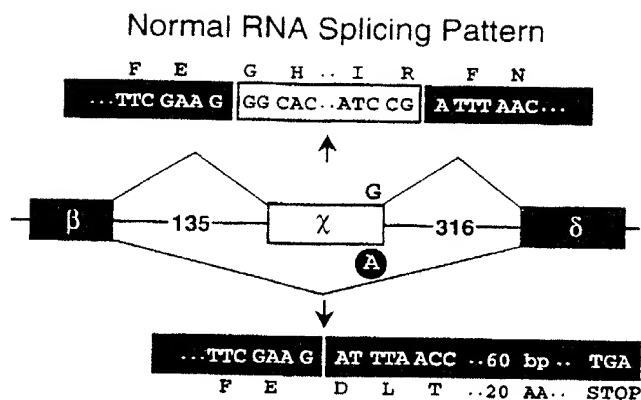


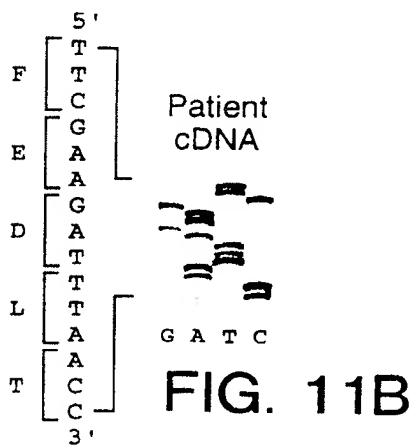
FIG. 10

RECTIFIED SHEET (RULE 91)



Patient RNA Splicing Pattern

FIG. 11A



RECTIFIED SHEET (RULE 91)



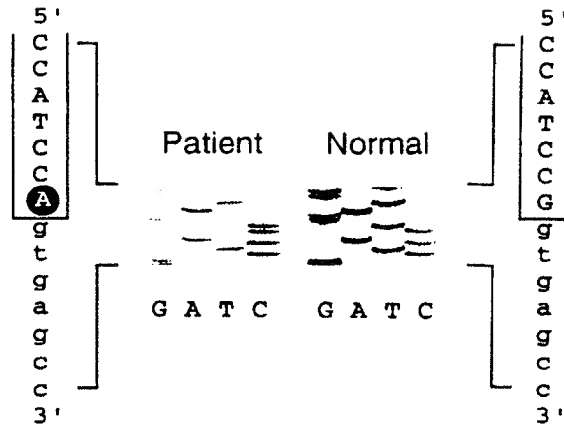


FIG. 11C

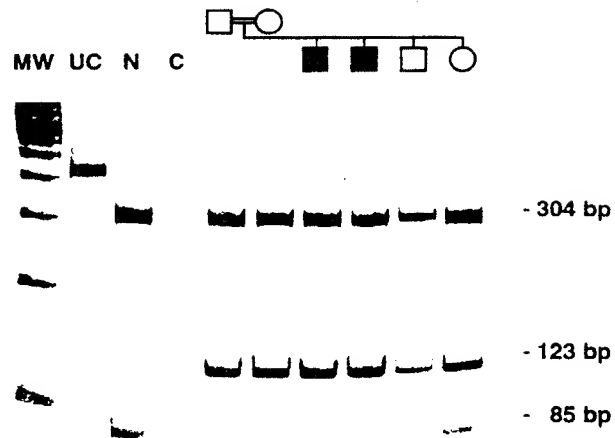


FIG. 11D

RECTIFIED SHEET (RULE 91)

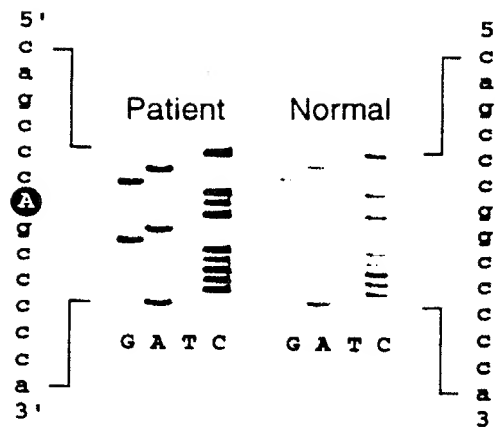


FIG. 12A

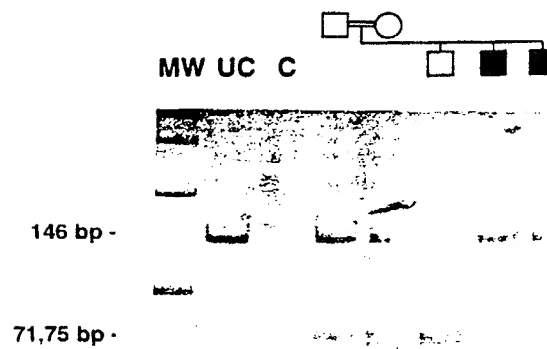


FIG. 12B

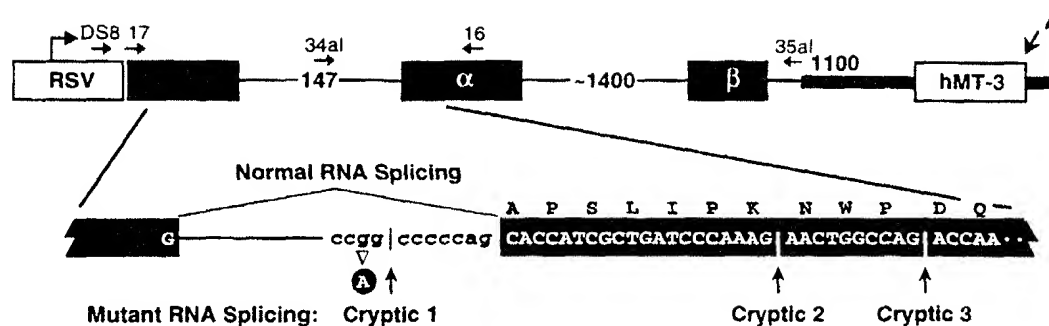


FIG. 12C

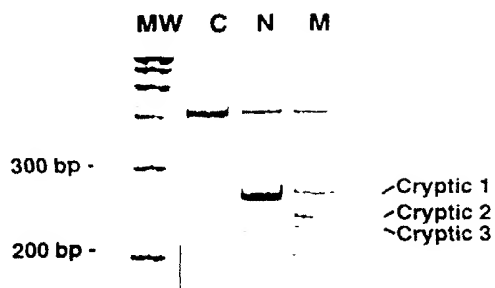


FIG. 12D

RECTIFIED SHEET (RULE 91)

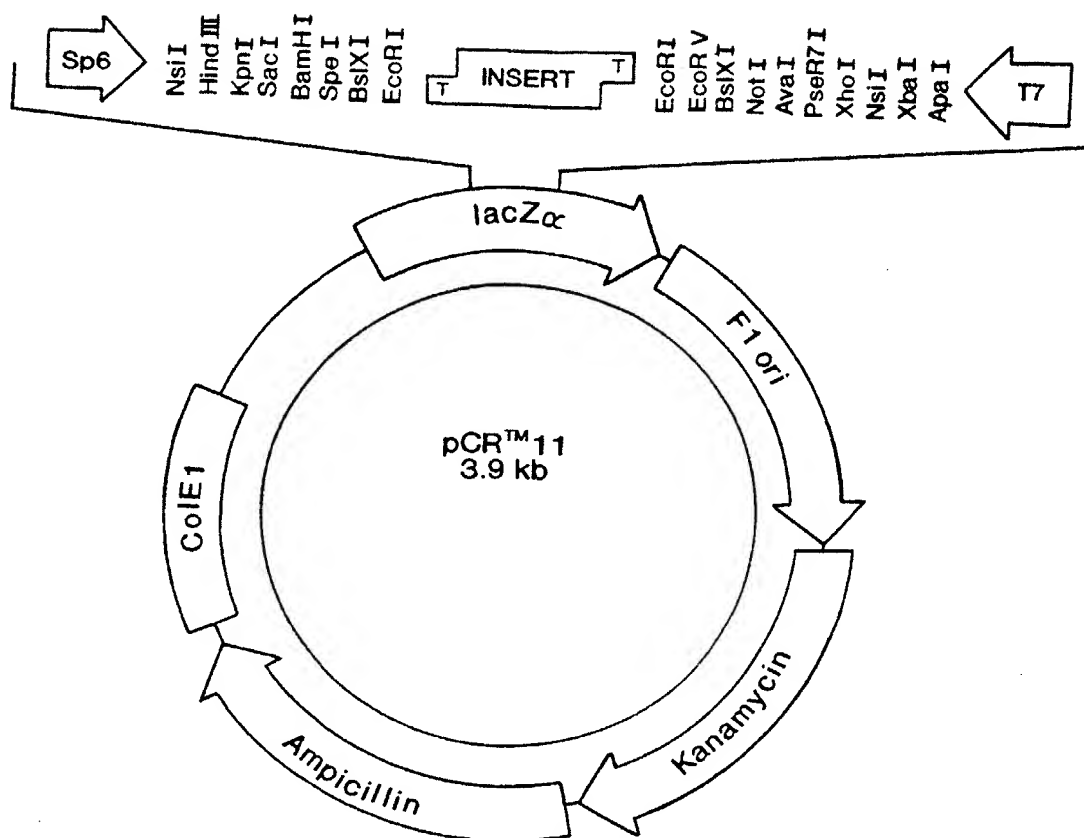


FIG. 13

RECTIFIED SHEET (RULE 91)

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US95/04463

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07H 15/12; C12N 15/00; C12Q 1/68

US CL : 536/23.1; 530/350+, 38; 800/2; 435/6

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 530/350+, 38; 800/2; 435/6

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS and Chemical Abstracts

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Proceedings of the National Academy of Sciences USA, Volume 89, issued July 1992, A. Virsolvy-Vergine et al, "Endosulfine, an Endogenous Peptide Ligand for the Sulfonylurea Receptor: Purification and Partial Characterization from Ovine Brain", pages 6629-6633, see entire document.	1-14, 17, 18, 35-43, 47 and 50
X	Advances in Genetics, Volume 24, issued 1987, G. Scangos et al, "Gene Transfer into Mice", pages 285-322, see especially pages 314 to 315.	15, 16 and 44-46
X	US, A, 4,965,188 (MULLIS ET AL) 23 October 1990, see entire document.	19-34, 48 and 49

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:	*T later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
*A document defining the general state of the art which is not considered to be of particular relevance	*X document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
*E earlier document published on or after the international filing date	*Y document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
*L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	*G document member of the same patent family
*O document referring to an oral disclosure, use, exhibition or other means	
*P document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search

08 JULY 1995

Date of mailing of the international search report

17 JUL 1995

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

Deborah Crouch, Ph.D.

Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)\*

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US95/04463

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

### Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.  
☐ No protest accompanied the payment of additional search fees.

Form PCT/ISA/210 (continuation of first sheet(1))(July 1992)\*

# INTERNATIONAL SEARCH REPORT

International application No.

PCT/US95/04463

## BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be examined, the appropriate additional examination fees must be paid.

Group I, claims 1-14, 17, 18, 35-43, 47 and 50, drawn to nucleic acids encoding a sulfonylurea receptor, expression vectors, host cells, protein and monoclonal antibodies.

Group II, claims 15, 16 and 44-46, drawn to transgenic non-human mammals.

Group III, claims 19-34, 48 and 49, drawn to methods to detect hyperinsulinemic hypoglycemia by amplification of DNA sequences in a sample.

The inventions listed as Groups I-III do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The inventions of groups I and II are distinct from each other as there is not a shared special technical feature. The production of transformed host cells of group I requires separate and distinct protocols from the production of transgenic non-human mammals of group II. The manipulation of isolated cells in culture as in group I requires considerations different from the manipulation of fertilized eggs as in group II. The cells of group I are not required to produce the transgenic non-human mammal of group II, and vice versa. The inventions of groups I and III do not share a special technical feature. The method of producing transformed host cells in group I requires separate and distinct protocols from the method of detection in group III. The manipulation of isolated cells in culture as in group I requires considerations different from the manipulation of DNA for amplification as in group III. The cells of group I are not required for the method of detection in group III and vice versa. The inventions of groups II and III do not share a special technical feature. The method for producing non-human transgenic mammals as in group II is separate and distinct from the method of detection as in group III. The manipulation of a fertilized egg requires considerations different from the manipulation of DNA for amplification. The mammal of group II is not required for the method of detection in group III, and vice versa. Accordingly the claims are not so linked by a special technical feature within the meaning of PCT Rule 13.2 so as to form a single inventive general concept.